



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**UTILIZING AN INTELLIGENT TUTORING SYSTEM IN
TACTICAL ACTION OFFICER SANDBOX**

by

Ryan W. McBride

June 2014

Thesis Advisor:
Co-Advisor:

Joseph Sullivan
Perry McDowell

**This thesis was performed at the MOVES Institute.
Approved for public release; distribution is unlimited**

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE June 2014	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE UTILIZING AN INTELLIGENT TUTORING SYSTEM IN TACTICAL ACTION OFFICER SANDBOX			5. FUNDING NUMBERS	
6. AUTHOR(S) Ryan W. McBride				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB protocol number NPS.2014.0030-IR-EP7-A.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (maximum 200 words) The Office of Naval Research Future Naval Capabilities, the Defense Technology Area Plan from 2005 and the Department of Defense Science and Technology Priorities for FY13-17 all share a focus on systems to promote warfighter performance. The goal of these systems is to improve effectiveness across a range of multiple missions, while providing training techniques that enhance fundamental cognitive abilities, institute techniques to reduce training time and training costs, maximize the training impact, as well as to provide tools and techniques to achieve routine and engaging scenario-based training. Intelligent tutoring systems (ITS) can address these goals. This work describes an ITS created for use in training naval surface warfare officers (SWOs) in air defense. A description of the learning theories applied in the development of the system, as well as a brief history of ITSs and the methodology that was utilized in developing this ITS is provided. A TCP/IP connection was established between this tutor and TAO Sandbox, a simulator utilized by Surface Warfare Officer School to train SWOs. Through a designed experiment, it was demonstrated that this ITS provides statistically significant training to its intended audience.				
14. SUBJECT TERMS Intelligent Tutoring System, ITS, Tactical Action Officer, TAO Sandbox, Cognitive Learning Theories			15. NUMBER OF PAGES 131	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

**UTILIZING AN INTELLIGENT TUTORING SYSTEM IN TACTICAL ACTION
OFFICER SANDBOX**

Ryan W. McBride
Lieutenant, United States Navy
B.S., United States Naval Academy, 2007

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN MODELING, VIRTUAL ENVIRONMENTS,
AND SIMULATION**

from the

**NAVAL POSTGRADUATE SCHOOL
June 2014**

Author: Ryan W. McBride

Approved by: Joseph Sullivan
Thesis Advisor

Perry McDowell
Co-Advisor

Peter J. Denning
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The Office of Naval Research Future Naval Capabilities, the Defense Technology Area Plan from 2005 and the Department of Defense Science and Technology Priorities for FY13-17 all share a focus on systems to promote warfighter performance. The goal of these systems is to improve effectiveness across a range of multiple missions, while providing training techniques that enhance fundamental cognitive abilities, institute techniques to reduce training time and training costs, maximize the training impact, as well as to provide tools and techniques to achieve routine and engaging scenario-based training.

Intelligent tutoring systems (ITS) can address these goals. This work describes an ITS created for use in training naval surface warfare officers (SWOs) in air defense. A description of the learning theories applied in the development of the system, as well as a brief history of ITSs and the methodology that was utilized in developing this ITS is provided.

A TCP/IP connection was established between this tutor and TAO Sandbox, a simulator utilized by Surface Warfare Officer School to train SWOs. Through a designed experiment, it was demonstrated that this ITS provides statistically significant training to its intended audience.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
A.	PROBLEM STATEMENT.....	1
B.	MOTIVATION.....	4
C.	APPROACH.....	4
II.	COGNITIVE THEORIES	5
A.	DREYFUS MODEL	6
B.	BLOOM'S TAXONOMY	8
C.	COGNITIVE LOAD THEORY.....	12
D.	INSTRUCTIONAL DESIGN THEORY.....	14
III.	COMPONENTS OF LEARNING.....	17
A.	FEEDBACK	17
B.	MOTIVATION.....	18
IV.	INTELLIGENT TUTORING SYSTEM	21
A.	HISTORY OF ITS.....	21
B.	PARTS OF ITS.....	22
C.	ITS AND SIMULATION IN THE NAVY	27
D.	POTENTIAL BENEFITS OF ITS AND SIMULATORS.....	28
V.	METHODOLOGY.....	31
A.	DESIGN OF EXPERIMENT	31
B.	SCOPE.....	34
C.	DEVELOPMENT OF ITS.....	34
VI.	TAO SANDBOX SCENARIO DESCRIPTION	39
A.	DETECT TO ENGAGE SEQUENCE.....	39
VII.	EXPERIMENT DATA AND ANALYTICAL RESULTS.....	47
A.	SUMMARY OF DATA	47
B.	ANALYSIS OF DATA	48
VIII.	DEMONSTRATION OF CONCEPT	57
IX.	LESSONS LEARNED.....	75
X.	CONCLUSIONS AND RECOMMENDATIONS.....	77
A.	FUTURE WORK.....	78
	APPENDIX A. ITS CODE	81
	APPENDIX B. RECORD OF CHANGES.....	91
	APPENDIX C. PRE-SCENARIO SURVEY	95
	APPENDIX D. POST-SCENARIO SURVEY	97
	APPENDIX E. NAVY SIMULATORS.....	99
	APPENDIX F. DATA EXTRACTED FROM TAO SANDBOX.....	105

LIST OF REFERENCES.....	109
INITIAL DISTRIBUTION LIST	113

LIST OF FIGURES

Figure 1.	Dreyfus model (from Sullivan & McDowell, 2013)	8
Figure 2.	Illustration of the basis of cognitive load theory (from University of South Alabama).....	13
Figure 3.	General ITS architecture (from Nwana, 1990, p. 257)	25
Figure 4.	Scenario 1	44
Figure 5.	Scenario 2	45
Figure 6.	Scenario 3	46
Figure 7.	NTA 1 Histogram of Experiment Group	49
Figure 8.	Wilcoxon rank sum of averages for all three runs.....	50
Figure 9.	Histogram of Control Deltas	51
Figure 10.	Histogram of Experiment Deltas.....	52
Figure 11.	Wilcoxon rank sum of deltas between runs one and three	53
Figure 12.	Means of deltas between runs 1 and 3 for control and experiment groups	54
Figure 13.	Standard Deviations of MOPs for control and experiment groups.....	54
Figure 14.	Breakdown of users in each group	55
Figure 15.	ITS introduction	61
Figure 16.	ITS suggesting turning RADAR on.....	62
Figure 17.	ITS suggesting HOOK action to user.....	63
Figure 18.	ITS suggesting evaluation of all contacts	64
Figure 19.	ITS suggesting user query unknown contact.....	65
Figure 20.	ITS providing feedback to user about successful query	66
Figure 21.	ITS providing feedback on results of reporting query results.....	67
Figure 22.	ITS suggesting the user warn unknown contact	68
Figure 23.	Simulator classifying unknown contact as hostile	69
Figure 24.	User initializing engagement of unknown contact with SAM.....	70
Figure 25.	ITS suggesting user report intentions of engagement	71
Figure 26.	ITS providing feedback of successful engagement assessment	72
Figure 27.	Example of interaction between simulator and tutor	73

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1.	ITS Authoring Tools with Strengths and Weaknesses (from Murray, 1999, p. 101)	27
Table 2.	Control Group Subjects 1-5	105
Table 3.	Control Group Subjects 6-10 and Averages	106
Table 4.	Experiment Group Subjects 1-5	107
Table 5.	Experiment Group Subjects 6-10 and Averages	108

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AI	artificial intelligence
AW	air warfare
ACAI	autonomous computer-aided instruction
BEWT	BFTT electronic warfare trainer
BFTT	battle force tactical trainer
CAI	computer-assisted instruction
CIC	combat information center
COVE	conning officer virtual environment
CG	guided missile cruiser
CVN	aircraft carrier, nuclear
DDG	guided missile destroyer
DETITWEE	detect, entry, tracking, identification, threat evaluation, weapons pairing, engagement, engagement assessment
FATS	firearms training simulator systems
FCA	fleet concentration areas
FFG	guided missile frigate
FMB	full mission bridge
FOV	field of view
GAO	government accountability office
GUI	graphical user interface
IMTS	intelligent maintenance training system
INCOFT	intelligent conduct of fire trainer
ITS	intelligent tutoring system
LCS	littoral combat ship
LHA	amphibious assault ship (general purpose)
LHD	amphibious assault ship (multipurpose)
LPD	amphibious transport dock
LSD	dock landing ship
MACH-III	maintenance aid computer for hawk – intelligent institutional instructor

MCM	mine countermeasures
MMTT	multi-mission tactical trainer
NCTE	navy continuous training environment
NSST	naval shiphandling simulator training
ONR	Office of Naval Research
PC	personal computer
PCO	prospective commanding officer
PORTS	PC based open architecture training system
PXO	prospective executive officer
QUEST	qualitative understanding of electrical system troubleshooting
S&T	science and technology
SOPHIE	sophisticated instructional environment
STCW	standards of training, certification and watchkeeping for seafarers
SWO	surface warfare officer
SWOS	surface warfare officer school
TAO	tactical action officer
TAO/ITS	tactical action officer/intelligent tutoring system
TRIO	trainer for radar intercept operations
TSSS	trainer simulator stimulator system
VR	virtual reality

ACKNOWLEDGMENTS

I would be remiss if I did not thank several individuals who were absolutely critical in the completion of this thesis.

First, I would like to thank LT Ethan Reber and LT Matt Georges. Both are NPS MOVES graduates. Had it not been for LT Reber's assistance in developing a thesis topic, I might still be trying to find a topic. Additionally, these two gentlemen were the sole Navy modeling and simulation assistance I received throughout the curriculum.

Next, I would very much like to thank Dr. Allen Munro. Dr. Munro graciously allowed me to utilize his simulator and spent many hours assisting me in navigating the TAO Sandbox, so that my ITS and thesis could be made possible.

I must thank my thesis advisors, CDR Joseph Sullivan and Perry McDowell. These two gentlemen spent countless hours advising me on my thesis and providing mentorship at a time in life when it was needed. NPS was not easy; however, these men assisted me in completing this adventure.

Additionally, I must thank my fellow cohort members. Each and every one of them added to my education in some fashion. Specifically, I must thank LTC David Parkes, USA. Dave is a true mentor and a great friend. As the lone Navy student in the curriculum, I was confused by the Marine and Army acronyms and systems. Dave acknowledged this and helped to educate me as I hope I educated him. Through this we developed a friendship that was paramount to my continued success at NPS.

Lastly, I must thank my wife, Shannon. Shannon is also a Navy SWO and is currently stationed in Norfolk, VA. Her support of my education and career were unwavering, despite the long distance and associated stresses. Without her, this endeavor would have been monotonous and incomplete. She is my

largest support system and the driving factor behind the completion of this thesis.
Behind every great man is a great woman. Shannon is a great woman.

I. INTRODUCTION

A. PROBLEM STATEMENT

The Office of Naval Research (ONR) has laid out many goals for the near future in its latest version of the Science and Technology (S&T) Plan that it released in 2011. One of those goals is to improve warfighter performance through “developing effective simulation-based training and creating computational cognitive models that accurately represent human training requirements” (p. 31). Furthermore, in the same document, ONR also states its desire to achieve the improvement of warfighter performance through manpower, personnel, training and education. Specifically, the object is to provide “training techniques to enhance fundamental information-processing abilities in young adults, institute techniques to shorten training time, reduce training costs and maximize training impact, as well as to provide tool and techniques to achieve ubiquitous, engaging, scenario-based training” (Office of Naval Research, 2011, p. 32).

It is widely known that the United States Navy and the Department of Defense, as a whole, use modeling and simulations for myriad tasks. One of those tasks is training. The Government Accountability Office (GAO) submitted a report to House and Senate Armed Forces Committee in June 2012, which reviewed the status of Navy training programs (United States Government Accountability Office, 2012). This review was directed as a rider for the National Defense Authorization Act for FY12 (H.R. Rep. No. 112-78 (2011)).

In this account, the GAO reported that over the past decade the Navy had drastically increased its emphasis on the use of utilizing modeling and simulation tools in its training, something the Navy calls “synthetic training”. Several examples of synthetic training include the submarine community conducting its pre-deployment training in shore-based simulators, the surface community conducting just over half of its training in simulators, and the aviation community

utilizing simulators for training new pilots as well as for continued proficiency (United States Government Accountability Office, 2012). Synthetic training allows the crews of multiple ships and aircraft to participate together in training events, practicing complex real-world operations that would be almost impossible to duplicate with actual platforms. In addition to simulating difficult tactical situations, synthetic trainers also reduce the wear and tear and operational expenses of using actual platforms for training (Greenert, 2012).

Additionally, in the Navy's Overarching Fleet Training Simulator Strategy, the Navy laid out several guiding principles that promote the use of simulators to the maximum extent possible. The guiding principles are:

1. "Effective training requires an efficient balance of live and synthetic approaches,"
2. "Simulator decisions are complex and require thoughtful and thorough analysis,"
3. "Train in port and validate at sea, or train on the ground and validate in the air, or train at home base and validate in the field,"
4. "Training simulators should be used to replace live training to the maximum extent possible where training effectiveness and operational readiness are not compromised,"
5. "Some live training events cannot or should not be replaced by a simulator,"
6. "If a skill or talent can be developed or refined, or if a proficiency can be effectively and efficiently maintained in a simulator, then these skills/talents/proficiencies should be developed/refined/maintained in a simulator,"
7. "If a qualification or certification can realistically and economically be accomplished in a simulator, do it in a simulator,"
8. "Simulator training objectives must be directly linked with specific Navy Mission Essential Tasks or individual personnel qualification standard requirements,"
9. "Simulators that are intended to interface with other simulators during Fleet Synthetic Training events must be compatible with the Navy Continuous Training Environment network,"

10. “Simulators that could conceivably be used for multi-platform or cross- platform mission area training should be designed with integration as a primary goal,”
11. “Simulators should provide the appropriate level of fidelity required to effectively and economically train to the specified task(s),”
12. “Simulator procurement needs to stay aligned with Fleet-wide technical innovation to deliver timely, cost effective solutions“
(United States Government Accountability Office, 2012, p. 14–15).

This thesis will examine Navy surface training. The Navy has several synthetic trainers that it uses for surface warfare training to include the Naval Shiphandling Simulator Training (NSST), Conning Officer Virtual Environment (COVE), Tactical Action Officer (TAO) Sandbox, Multi-Mission Tactical Trainer (MMTT), and Battle Force Tactical Trainer (BFTT). These synthetic trainers give the surface fleet the ability to conduct just over half of their training synthetically, but the mix of actual training varies by ship type and mission area. For example, some of the oldest ships do not have the type of computer infrastructure needed to support extensive synthetic training and require special technical assistance during Fleet Synthetic Training events. Conversely, simulators for one of the newest ships, the Littoral Combat Ship, are intended to provide crews with full certification prior to deployment, much like submarine crews (United States Government Accountability Office, 2012).

I believe that an intelligent tutoring system (ITS) could provide a solution to ONR’s goal of improving naval training. An ITS is a computer program that is designed to incorporate techniques from the artificial intelligence (AI) community to provide computerized tutors that know the subjects they are teaching, who they are teaching, and how best to teach them. An ITS can meet the specific objectives laid out in ONR’s S&T plan, which is to enhance the current training in a manner that it can save time and money as well as instruct young adults, like myself. To demonstrate this, the goal of my thesis is to create an ITS for a naval task and to conduct an experiment to determine whether the use of this ITS improves performance.

An ITS can be instituted in various mediums. For this particular thesis, I have linked one I created to the TAO Sandbox, which Surface Warfare Officer School (SWOS) currently uses as a scenario-based training tool in the Department Head curriculum. Specifically I will be answering three questions:

1. What learning theories must be considered when developing an expert pedagogical model which will provide feedback to students?
2. Does the intelligent tutoring system function in a manner that it could be utilized to adequately train a Surface Warfare Officer (SWO) in air defense?
3. What sequencing of events must be considered when responding to an air defense engagement, particularly within a detect to engage sequence?

B. MOTIVATION

As a SWO, my motivation lies in increasing the synthetic capability for the surface fleet. Synthetic training is utilized extensively throughout the training pipeline and throughout pre-deployment workups, which is directed by the Navy's Overarching Fleet Training Simulator Strategy. An effective ITS would help the Navy meet several of the guidelines established in the Navy's Overarching Fleet Training Simulator Strategy, help achieve goals set forth in the ONR S&T Plan, as well as work toward Future Naval Capabilities of the United States Navy.

C. APPROACH

The approach of this thesis is not a novel approach, but rather, a standard approach that is accepted throughout the research community. First I identified a problem, which is to develop and test an intelligent tutoring system that can help increase the Navy's ability to utilizing synthetic training. I then developed an experiment that could be utilized in proving, or disproving, my hypothesis. In order to execute the experiment, I developed three scenarios that would adequately represent an air warfare (AW) environment. Furthermore, I had to find a simulator, which is being utilized for naval training and adapt it to either embed or communicate with an ITS I developed which is based on the detect to engage (DTE) sequence used in many warfare areas.

II. COGNITIVE THEORIES

Pavlov, Frolov, Roessler, Brogden, Hull, Bugelski, Coyer, Miller, Skinner, Gagné, Bloom, Dreyfus and many others have all attempted to determine why people learn, what learning is, and what motivates individuals to learn. Most of them determined that learning is not a standard process for any one person or species. Some learning theories that are important to this study are the Dreyfus model, Bloom's taxonomy, cognitive load theory, and instructional design theory. Of the plethora of cognitive theories these are the theories I deemed to have the most influence on developing an ITS for an adult. I looked at theories that discussed the types of users, the types of ways to present materials and the types of ways to get through to students. This is due to the fact that an ITS is a system that knows what to teach, how to teach it, and who they are teaching it to.

The Dreyfus model will be utilized to determine if there is a user group that is more apt to be aided by the ITS than another. Bloom's taxonomy can be thought of as a framework for categorizing educational goals and given the major goal of an ITS is education, it will be helpful to categorize the goals being presented by the ITS. Cognitive load theory uses evolutionary theory to consider human cognitive architecture and uses that architecture to devise novel, instructional procedures. Instructional design theory offers a template for developing and delivering instruction, which is instrumental in teaching. If an instructional tool is not developed and delivered in a proper manner, it may not provide adequate instruction.

I chose these particular theories because they address adult learning and collectively they help address the creation of an ITS. Additionally, they are theories that I am more familiar with, from the instruction received throughout the MOVES curriculum. Utilizing these theories, I should be able to demonstrate that the ITS is a novel instructional tool that can assist in teaching adults. I will utilize

these four theories in developing the ITS and I postulate that an ITS could adequately instruct the right audience in a manner which will be effective and lasting.

A. DREYFUS MODEL

The Dreyfus and Dreyfus (1986) five-stage model of skill acquisition characterizes performance levels through which individuals progress as they gain skill and proficiency in cognitively complex domains. The model has been applied to training and instruction with domains such as combat aviation, nursing, industrial accounting, psychotherapy, and chess (Phillips, Shafer, Ross, Cox, & Shadrack, 2006). The stages of the Dreyfus model are:

Stage 1: Novice. Novices have limited or no experience in situations characteristic of their domain. They exhibit rigid adherence to rules they have been taught, or plans they have been given. They have little situational perception, and they lack the basic domain knowledge needed to perform analysis.

Stage 2: Advance Beginner. Advanced beginners have enough domain experience that their performance is marginally acceptable. They have a sufficient knowledge base with which to analyze a situation. At this stage they are able to recognize recurring, meaningful “aspects” of situations global characteristics identifiable only through prior experience where the prior experience provides a comparison case for the current situation. Their knowledge base regarding aspects and attributes of situations enable them to develop their own guidelines for action. However, all components of the situation tend to be treated as independent pieces and as equal in importance, rather than differentially weighted based on the circumstances and goals.

Stage 3: Competent. At the competent level, performers have mental models that they can apply to the new situations. This stage is

marked by the ability to envision and predict how a situation is likely to play out, which guides the formulation, prioritization, and management of longer-term goals. Competent performers are very planful, where advance beginners are more reactive. However, competent individuals tend to adhere to the plan as the situation plays out, even when circumstances change. They have difficulty adapting their plan to address new situational demands.

Stage 4: Proficient. Proficient individuals' performance shifts from being guided by the plan to being responsive to the situation. They see the situation as an inseparable whole rather than independent attributes; they have the ability to recognize meaningful patterns of cues without breaking them down into their component parts for analysis. As such, they are able to intuitively assess what is happening and what is most critical for achieving success. They shift their assessment of the situation as it evolves and changes, and they can adjust their course of action accordingly. However, while their situation assessment is recognitional and intuitive, they still perform deliberate analysis when making decision and devising or adjusting a course of action.

Stage 5: Expert. Expert performance is marked by a shift to recognitional decision-making. Experts intuitively assess the situation and also intuitively recognize a suitable course of action that will accomplish their goals. They have a substantial base of experience from which to operate. Their mental models are broad, deep, and elaborate. They are able to make fine discriminations between perceptual cues and can diagnose and assess situations that confuse or stump their less experienced peers. Experts also have a wide range of routines and tactics for getting things done (Phillips, Shafer, Ross, Cox, & Shadrick, 2006).

This is often illustrated as such:

The Dreyfus Model



Figure 1. Dreyfus model (from Sullivan & McDowell, 2013)

This model will become important when discussing the type of user interfacing with the ITS. In preliminary research it has been shown that an ITS is good for the beginner but not quite as good for an expert. This has to do with presentation of material and feedback usually in a quantity that the expert feels is unnecessary. The design of experiment was purposefully setup to grab from this entire spectrum to demonstrate, on a small scale, whether this will hold true for this particular ITS.

B. BLOOM'S TAXONOMY

In 1956, Benjamin Bloom, with collaborators Max Englehart, Edward Furst, Walter Hill, and David Krathwohl, published a framework for categorizing educational goals: *Taxonomy of Educational Objectives*. This is known today as

Bloom's taxonomy. This framework has been applied by generations of K-12 teachers and college instructors in their teaching. The framework elaborated by Bloom and his collaborators consisted of six major categories: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. The categories after Knowledge were presented as "skills and abilities," with the understanding that knowledge was the necessary precondition for putting these skills and abilities into practice (Armstrong, 2014). The six major categories are described as:

1. **Knowledge** "involves the recall of specifics and universals, the recall of methods and processes, or the recall of a pattern, structure, or setting."
2. **Comprehension** "refers to a type of understanding or apprehension such that the individual knows what is being communicated and can make use of the material or idea being communicated without necessarily relating it to other material or seeing its fullest implications."
3. **Application** refers to the "use of abstractions in particular and concrete situations."
4. **Analysis** represents the "breakdown of a communication into its constituent elements or parts such that the relative hierarchy of ideas is made clear and/or the relations between ideas expressed are made explicit."
5. **Synthesis** involves the "putting together of elements and parts so as to form a whole."
6. **Evaluation** engenders "judgments about the value of material and methods for given purposes." (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956, p. 201-207)

In 2001, a group cognitive psychologists, curriculum theorists and instructional researchers, and testing and assessment specialists revised this theory. The revised theory maintained six categories but broke knowledge into its own taxonomy. According to Armstrong, the new theory "underscores this dynamism, using verbs and gerunds to label their categories and subcategories (rather than the nouns of the original taxonomy). These "action words" describe

the cognitive processes by which thinkers encounter and work with knowledge” (Armstrong, 2014).

The revised taxonomy takes shape as:

1. Remember
 - a) Recognizing
 - b) Recalling
2. Understand
 - a) Interpreting
 - b) Exemplifying
 - c) Classifying
 - d) Summarizing
 - e) Inferring
 - f) Comparing
 - g) Explaining
3. Apply
 - a) Executing
 - b) Implementing
4. Analyze
 - a) Differentiating
 - b) Organizing
 - c) Attributing
5. Evaluate
 - a) Checking
 - b) Critiquing

6. Create
 - a) Generating
 - b) Planning
 - c) Producing

They also broke knowledge out on its own in this manner:

1. Factual Knowledge
 - a) Knowledge of terminology
 - b) Knowledge of specific details and elements
2. Conceptual Knowledge
 - a) Knowledge of classifications and categories
 - b) Knowledge of principles and generalizations
 - c) Knowledge of theories, models, and structures
3. Procedural Knowledge
 - a) Knowledge of subject-specific skills and algorithms
 - b) Knowledge of subject-specific techniques and methods
 - c) Knowledge of criteria for determining when to use appropriate procedures
4. Metacognitive Knowledge
 - a) Strategic Knowledge
 - b) Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge
 - c) Self-knowledge

Armstrong suggests several reasons as to why we use Bloom's taxonomy:

1. Objectives (learning goals) are important to establish in a pedagogical interchange so that teachers and students alike understand the purpose of that interchange.
2. Teachers can benefit from using frameworks to organize objectives because
3. Organizing objectives helps to clarify objectives for themselves and for students.
4. Having an organized set of objectives helps teachers to:
 - a) “plan and deliver appropriate instruction”;
 - b) “design valid assessment tasks and strategies”; and
 - c) “ensure that instruction and assessment are aligned with the objectives” (Armstrong, 2014).

Teachers are not the only benefactors of this taxonomy. Bloom’s taxonomy can extend to an ITS, as it strives to perform the sole duty of teachers, which is instruction.

C. COGNITIVE LOAD THEORY

Cognitive load theory uses evolutionary theory to consider human cognitive architecture and uses that architecture to devise novel, instructional procedures. The theory assumes that knowledge can be divided into biologically primary knowledge that we have evolved to acquire and biologically secondary knowledge that is important for cultural reasons. Secondary knowledge, unlike primary knowledge, is the subject of instruction (Sweller, 2011).

Cognitive load theory is important because it discusses how working memory is the key to achieving efficient learning. Working memory is the limited capacity part of human memory that allows us to temporarily store information, manipulate it and use it in our thinking (Baddeley & Hitch, 2011; Pearson Education, Inc., 2014). For a good description of working memory and how we use it, see (Pearson Education, Inc., 2014). Cognitive load theory suggests that the trainee’s ability to use working memory changes as expertise develops and

the training systems should be tailored around trainee's current working memory capacity. This can best be illustrated as such:



Figure 2. Illustration of the basis of cognitive load theory
(from University of South Alabama)

The idea behind this theory is that we cannot change the intrinsic load; however, we can manipulate the germane and extraneous load. The intrinsic load depends completely on the complexity of the to-be-learned content. It is the memory required by the thinking task at a given time and measures the amount of the working memory in use due to the interactivity of the amount of information being processed (University of South Alabama). This cannot be modified by instructional design. An example of this is solving a simple addition problem versus solving a differential equation. The differential equation will be more difficult than the addition problem no matter how it is presented to the user.

Germane load is the load that helps building new complex schema in a successive manner helping the learner to move from novice to expert. It is a self-effort to learn, and memorize information learned (University of South Alabama). An example of this would be highlighting words in text to demonstrate what words might be essential to know.

Extraneous cognitive load results from the techniques in which the to-be-learned information is presented. This load does not contribute to learning and can be modified by instructional design. It can be changed in a variety of ways, by enhancing the organization, chunking, and presentation techniques of to-be-learned information. Other ways are by using adjunct aids, and providing specific

learning instructions (University of South Alabama). Examples of extraneous cognitive load are multiple sources of information such as multiple sources of information, long complex explanations, or extra sounds.

Cognitive load theory plays an important role in developing an ITS as it suggests that in order to accurately and successfully teach someone we must be present the material in manner in which we reduce the extraneous cognitive load and maximize the germane load. In applying this theory, I developed scenarios that were limited in scope and would allow the DTE sequence to be taught in a manner which focuses on the air defense task with very little extraneous distractions.

D. INSTRUCTIONAL DESIGN THEORY

When developing the expert system of the ITS it is important to design the system in a manner consistent with the instructional design process. An expert system of an ITS will be fully explained in the ITS chapter; however, for now it will be sufficient to understand that it is the process that an expert would take in order to complete the task being tutored. Instructors have used the instructional design process for decades and it has been well tested in many venues, so following it correctly gives confidence that the product will provide useful instruction. The process has aided in the development of many of the teaching and learning processes that we use to today in our school systems, raising kids, and job training (Gagné & Briggs, 1979).

In this process Gagné suggests that there are certain assumptions that must be made in instruction design. These assumptions are that instructional design:

1. is aimed at “aiding the learning of the individual”;
2. it has phases that are “immediate and long term”;
3. “systematically designed instruction can greatly affect individual human development”;

4. "instructional design must be conducted by means of a system approach";
5. that "designed instruction must be based on knowledge of how human beings learn" (Gagné & Briggs, 1979, p. 4-5).

In further explaining how learning occurs, Gagné breaks his theory down into nine functional events of instruction. These are:

1. "Gain attention. Present a good problem, a new situation, use a multimedia advertisement, and ask questions. This helps to ground the lesson, and to motivate."
2. "Describe the goal. State what students will be able to accomplish and how they will be able to use the knowledge, give a demonstration if appropriate. This allows students to frame information."
3. "Stimulate recall of prior knowledge. Remind the student of prior knowledge relevant to the current lesson (facts, rules, procedures or skills). Show how knowledge is connected, provide the student with a framework that helps learning and remembering. Tests can be included."
4. "Present the material to be learned. This can be done with text, graphics, simulations, figures, pictures, sound, etc. Chunk information (avoid memory overload, recall information)."
5. Provide guidance for learning. Presentation of content is different from instructions on how to learn. Use of different channel (e.g. side-boxes)
6. "Elicit performance "practice". Let the learner do something with the newly acquired behavior, practice skills or apply knowledge."
7. "Provide informative feedback. Show correctness of the trainee's response, analyze learner's behavior, and maybe present a good (step-by-step) solution of the problem."
8. "Assess performance test, if the lesson has been learned. Also give sometimes general progress information."
9. "Enhance retention and transfer. Inform the learners about similar problem situations, provide additional practice. Put the learner in a transfer situation. Maybe let the learner review the lesson" (EduTech Wiki, 2007).

At the end of the day, Gagné's most essential ingredients of teaching are:

1. Presenting the knowledge or demonstrating the skill
2. Providing practice with feedback
3. Providing learner guidance

The development of my ITS expert model lends heavily to Gagné's instructional design theory. This was intentional as Gagné derived his instructional design theory from years of work in military training, the topic of interest for this thesis. I used the nine events to help develop the ITS; paying particular attention to how it was presented to the user as well as how it provided feedback and learner guidance.

III. COMPONENTS OF LEARNING

There are many components to learning. When discussing human factors within simulators throughout the curriculum, I understood motivation and feedback to be critical. This understanding is based upon the many discussions in human factors that I experienced while completing the MOVES curriculum. These discussions often centered on the ability to capture and maintain the user's attention as well as how to properly train the user, utilizing simulations and virtual environments. Capturing and maintaining a user's attention continually referenced motivation of the user as well as feedback to the user. As such, I focused on these two components of learning when it came to designing my ITS.

A. FEEDBACK

Feedback is critical in teaching students. Timely feedback is a critical feature of cognitive tutors that lead to cognitive and motivational benefits. So what is the right timing of feedback? Most studies show that immediate feedback is better than delayed or no feedback. In a parametric study done by Corbett and Anderson in 1991 utilizing the LISP tutor, it was shown that the immediacy of feedback leads to dramatic reductions in the learning time need to reach the same level of post-training performance. Learning time was shown to be three times longer in the most delayed feedback conditions than in the most immediate (Koedinger, Anderson, Hadley, & Mark, 1997). Further research has shown that there are other reasons for desiring immediacy of feedback to include psychological evidence that feedback on an error is effective to the degree that it is given in close proximity to the error, immediate feedback makes learning more efficient because it avoids long stretches of time where a student stumbles through an incorrect solution, and it tends to avoid the extreme frustration that builds up as the student struggles in a state of error (Anderson, Boyle, Corbett, & Lewis, 1990).

While immediate feedback is suggested by this research, there are a number of problems with utilizing immediate feedback. The feedback must be designed in a manner that forces a student to think, not just copy the correct answer. If given a few more seconds or minutes a user might have noticed the error themselves; self-correction is preferable because people tend to remember better what they generate themselves. Additionally students find immediate correction annoying, which is particularly true for more experienced students (Anderson, Boyle, Corbett, & Lewis, 1990).

In the development of an ITS within COVE by a team at Stanford University, it was determined that as a majority of feedback came after an action was taken, it was imperative to allow the student to make a mistake first (Wong, Kirchenbaum, & Peters, 2010). Similarly, when considering feedback, I determined that I would provide immediate feedback; however, I would allow the users time to conduct the action prior to correcting them.

B. MOTIVATION

When considering motivation, it was critical to consider the individual being taught. As with feedback, I wanted to provide the right motivation for the individuals. Maintaining motivation is as critical as establishing motivation. One factor of motivation in developing an ITS is time. Students are unlikely to patiently wait as a system takes time computing solutions or trying to figure out what the student is doing (Anderson, Boyle, Corbett, & Lewis, 1990). In order to not bog down the computer with computations, I focused on the DTE sequence, MOPs, and minimized computations required of the ITS. Measures of performance (MOPs) are the measurements of a system's performance typically expressed as speed, payload, range, time-on-station, frequency, or other distinctly quantifiable performance features (Defense Acquisition University, 2010). I also increased the size of available RAM and changed IDEs to provide a smooth interaction between the simulator and the ITS.

In the previous section I spoke about the timing of feedback. This feedback also plays into motivation. It allows students to know right away if they are making progress with challenging tasks. Further, if feedback minimizes the severity of errors, students may not feel the social stigma associated with making an error in class or potentially in front of a commanding officer or peers. Additionally by keeping students engaged in successful problem solving by utilizing feedback it can reduce student frustration and provide for a sense of accomplishment (Koedinger, Anderson, Hadley, & Mark, 1997).

Another motivation that I considered was the interest level of the student participating. When developing the scenarios, I attempted to make them complex enough that all users could be interested in the task at hand while trying to not overwhelm the user. This balance appears to have worked as 18 of 20 participants felt the scenario was realistic and only a handful felt the need to speed up the problem within the TAO Sandbox.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. INTELLIGENT TUTORING SYSTEM

An intelligent tutoring system (ITS) is a complex computer program that manages the various heterogeneous types of knowledge, ranging from domain to pedagogical knowledge.

A. HISTORY OF ITS

ITSs sometimes referred to as intelligent computer-assisted instruction (ICAI), have existed since the early 1980s. In this time, there have been many intelligent tutors designed for myriad purposes, mostly in science, technology, engineering, and mathematics (STEM) topic areas. The precursor to ITS was computer-assisted instruction (CAI), which came about in the mid-1960s. Some of the first CAI were developed by Leonard Uhr and his collaborators, when they implemented a series of systems which generated a protocol for arithmetic and vocabulary-recall. Shortly after more systems followed that were implemented to provide “drill and practice” in arithmetic and that would select problems based upon the users performance (Sleeman & Brown, 1982). Decades of years to define time spans are plural (with an s).

Many other systems have followed since and have continually furthered the field; a field whose original goal of extending the domains of applicability, power and accuracy of adaptive CAIs. But to do so, the system must examine more than just an answer, it must examine the methodology the student used to arrive at said answer and extrapolate what the student was thinking. This is where the cognitive theories truly took hold in the ITS field. However, all of the early systems and arguably the present day systems have some shortcomings. Some of the admitted shortcomings of earlier systems, noted in the book *Intelligent Tutoring Systems* by Sleeman and Brown (1982) include:

1. “The instructional material produce in response to a student’s query or mistake is often at the wrong level of detail, as the system assumes too much or too little student knowledge.”

2. "The system assumes a particular conceptualization of the domain, thereby coercing a student's performance into its own conceptual framework. None of the systems can discover, and work within, the student's own (idiosyncratic) conceptualization to diagnose his 'mind bugs' within that framework."
3. "The tutoring and critiquing strategies used by these systems are excessively ad hoc reflecting unprincipled intuitions about how to control their behavior. Discovering consistent principles would be facilitated by constructing better theories of learning and mislearning (negative transfer) "
4. "User interaction is still too restrictive, limiting the student's expressiveness and thereby limiting the ability of the tutor's diagnostic mechanisms" (p. 240-241).

B. PARTS OF ITS

In the early days of ITS development, there were three basic components of an ITS. These components were the expert module, the student module, and the tutoring module. Follow on research has lent itself to including a fourth module, which is referred to as the user interface module (Nwana, 1990; Freedman, 2000)

The expert model, also known as the domain or cognitive model, are the facts and rules of the particular domain being taught to the student (i.e., the knowledge of the experts). The expert model routinely performs two functions. The first function is that it is the source of knowledge presented to the students. The second function is that it provides a standard for evaluating the student's performance (Nwana, 1990).

The student model refers to the dynamic representation of the emerging knowledge and skill of the student. No intelligent tutoring can take place without an understanding of the student. Thus, along with the idea of explicitly representing the knowledge to be communicated, came the idea of doing likewise with the student, in the form of a student model (Nwana, 1990). This model is possibly the most complicated model to build as this is where the tutor attempts to determine what the student is thinking when they make decisions and

why certain decisions were made. Nwana states in the Intelligent Tutoring Systems paper that, ideally, “this model should include all those aspects of the student's behavior and knowledge that have possible repercussions on his/her performance and learning. However, the task of constructing such a complete model is not only non-trivial but, probably, impossible, especially considering that the communication channel, which is usually the keyboard, is so restrictive” (1990, p.260). Nwana continues by generally classifying the uses of a student model as such:

1. “Corrective: to help eradicate bugs in the student's knowledge.”
2. “Elaborative: to help correct 'incomplete' student knowledge.”
3. “Strategic: to help initiate significant changes in the tutorial strategy other than the tactical decisions of 1 and 2 above.”
4. “Diagnostic: to help diagnose bugs in the student's knowledge.”
5. “Predictive: to help determine the student's likely response to tutorial actions.”
6. “Evaluative: to help assess the student or the ITS” (1990, p. 260).

The tutoring model, or teaching model, is the part of the ITS that designs and regulates instructional interactions with the student. This is the source of all pedagogical interventions. Nwana suggests “the order and manner in which topics are treated can produce very different learning experiences.” (1990, p. 261) In tutoring, it is often more effective to allow the student to attempt a scenario or problem on their own and even allow them to get stuck for a while before the tutor will interrupt. As a result of this knowledge, there are several different manners in which a tutoring model can act. This ranges from a monitoring mode to guided-discovery learning systems and all manner of tutoring in between to include the midway approach of a mixed-initiative system. In a monitoring mode the system monitors the students every action, adapting the ITS actions to the student responses but never truly relinquishing control (Nwana, 1990). The guided-discover learning mode allows the student full control and can

only intervene by manipulating the environment or scenario. The middle approach or the mixed-initiative system allows the student and the system to share control and interaction occurs through a question and answer forum. The student and ITS can ask each other questions and the appropriate answer or reaction is given (Nwana, 1990). The many ways in which a tutoring model can interact demonstrates the art form that is instructing. There is no right way other than the way in which the student learns the material. Further complicating the implementation of an ITS is that many of the subject areas that they are designed for are art forms themselves. The importance of the tutor model is to assist the student, while not destroying the sense of personal motivation the student maintains or the students' sense of discovery. The human brain is a remarkable organ and the thirst for knowledge must be encouraged.

The user interface model is the communicating component of the ITS, which controls interaction between the student and the system. Previously, this was not considered a component of an ITS but as the field has furthered, it has been determined that this model can make or break the ITS, no matter how good the student model or expert model may be. There are two reasons why this model is so important according to Nwana, "First, when the ITS presents a topic, the interface can enhance or diminish the presentation. Since the interface is the final form in which the ITS presents itself, qualities such as ease of use and attractiveness could be crucial to the student's acceptance of the system. Secondly, progress in media technology is increasingly providing more and more sophisticated tools whose communicative power heavily influences ITS design"(1990, p. 262).

The general ITS architecture looks like:

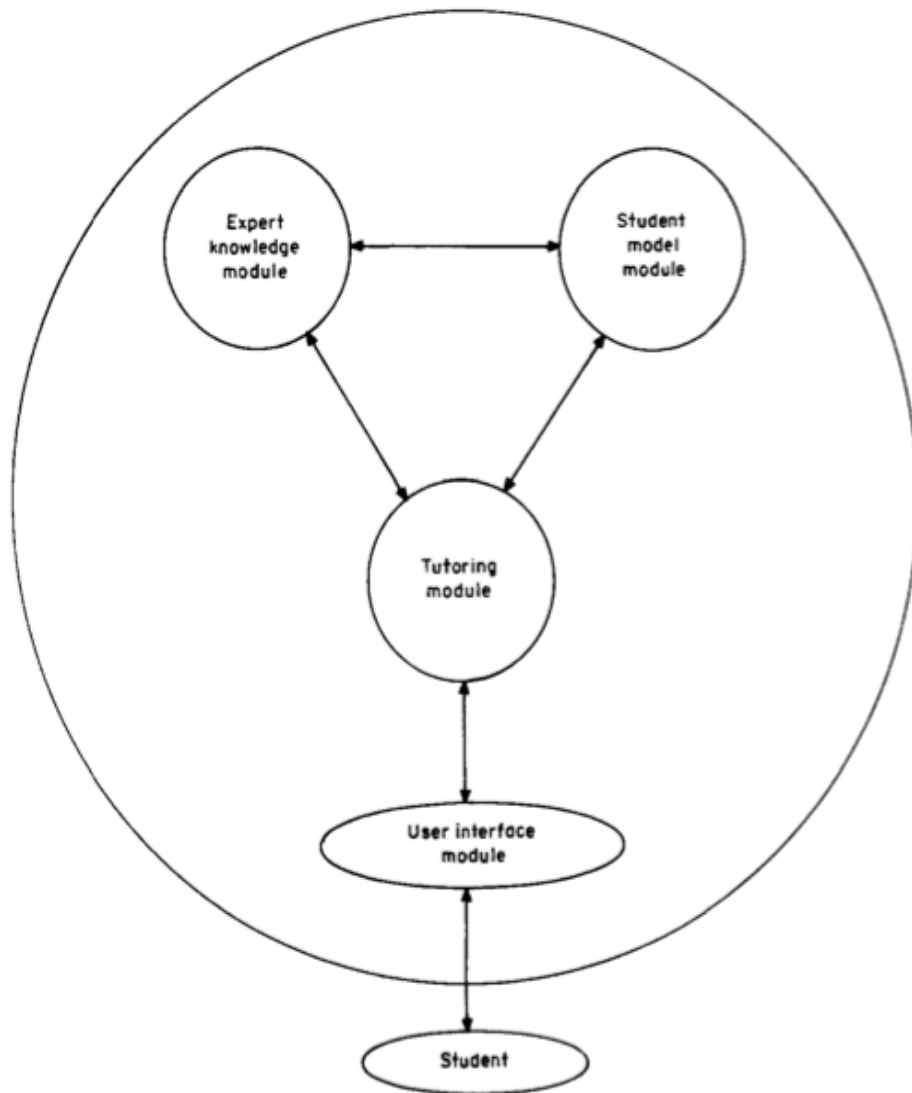


Figure 3. General ITS architecture (from Nwana, 1990, p. 257)

As suggested earlier, tutoring is an art form and I would postulate that developing an ITS is also an art form. Simply, there is no right way to code an ITS. In fact there are numerous ways and numerous tools to assist in this. The table below shows the many types of ITS authoring tools that are available as well as the strengths and limitations of each.

Category	Strengths	Limits	Variations
Curriculum Sequencing and Planning	Rules, constraints, or strategies for sequencing courses, modules, presentations	Low fidelity from student's perspective; shallow skill representation	Whether sequencing rules are fixed or authorable; scaffolding of the authoring process
Tutoring Strategies	Micro-level tutoring strategies; sophisticated set of instructional primitives; multiple tutoring strategies	(same as above for most systems)	Strategy representation methods; source of instructional expertise
Device Simulation and Equipment Training	Authoring and tutoring matched to device component identification, operation, and troubleshooting	Limited instructional strategies; limited student modeling; mostly for procedural skills	Fidelity of the simulation; ease of authoring
Domain Expert System	Runnable (deeper) modal of domain expertise; fine grained student diagnosis and modeling; buggy and novice rules included	Building the expert system is difficult; limited to procedural and problem solving expertise; limited instructional strategies	Cognitive vs. performance models of expertise
Multiple Knowledge Types	Differential pre-defined knowledge representation and instructional methods for facts, concepts and procedures, etc.	Limited to relatively simple fact, concepts, and procedures; pre-defined tutoring strategies	Inclusion of intelligent curriculum sequencing; types of knowledge/tasks supported
Special Purpose	Template-based systems provide strong authoring guidance; fixed design or pedagogical principles can be enforced	Each tool limited to a specific type of tutor; inflexibility of representation and pedagogy	Degree of flexibility

Intelligent/Adaptive Hypermedia	WWW has accessibility and UI uniformity; adaptive selection and annotation of hyperlinks	Limited interactivity; limited student model bandwidth	Macro vs. micro level focus; degree of interactivity
---------------------------------	--	--	--

Table 1. ITS Authoring Tools with Strengths and Weaknesses
(from Murray, 1999, p. 101)

As recently as 2005, there was a push to standardize the ITS. Since an ITS is a conglomeration of artificial intelligence, the subject matter in which you are teaching, cognitive studies, psychology, computer science, simulators (as technology increases, more and more training is being conducted on simulators), and a plethora of other fields it has proved to be a difficult process and has yet to be completed.

C. ITS AND SIMULATION IN THE NAVY

In the military there has been a huge push for simulators that can be used for training. This is largely due to the recent reduction in training budgets. The throughput of a simulator as compared to real systems can be much higher. As the push for simulators and the realization of technology continues, so does the popularity of ITSs and the potential for what they can accomplish. Bob Pokorny, a long time researcher in ITSs, wrote an issue of *Military Simulation and Training* that “Gen. Martin E. Dempsey, Chairman of the Joint Chiefs of Staff, noted a need for the military to train with realistic, complex, and unpredictable scenarios, followed by reviews that feed subsequent instruction.” (Pokorny, 2012, p. 21) Pokorny also noted that this same vision and recommended experiential instruction in which remediation adapts to each trainee’s needs is embodied in the Army Learning Concept 2015 and suggests that this “technology provides the immersive, complex practice environments in which assessment of trainees’ performance guides individually adapted remediation” (Pokorny, 2012, p. 22).

The push for simulations and ITS can be seen as early as 1989, when there were several ICAI that were in use in the military. These include

Sophisticated Instructional Environment (SOPHIE), STEAMER, Qualitative Understanding of Electrical System Troubleshooting (QUEST), Intelligent Maintenance Training System (IMTS), Maintenance Aid Computer for Hawk – Intelligent Institutional Instructor (MACH-III), Trainer for Radar Intercept Operations (TRIO), and Intelligent Conduct of Fire Trainer (INCOFT) (Salgado-Zapata, 1989). Since then there has been an increase in simulator use and in the past few years a huge push towards developing ITSs for those simulators. See Appendix E for a description of the simulators listed above, as well as a description of current simulators being utilized to train SWOs.

D. POTENTIAL BENEFITS OF ITS AND SIMULATORS

There are many benefits to an ITS and simulators in general. The primary benefit in our current fiscally constrained environment is that these systems have the potential to save money. Currently these systems are expensive but this is due to the domain-specific nature of ITSs.

They can improve training. There are several ITSs that have been developed that have shown that an ITS is viable educational tool. One example is PAT, which was designed for an algebra curriculum and employed in the Pittsburgh school district during the 1993-94 school year. PAT showed that on average, the 470 students who utilized the system outperformed students in comparison classes by 15% on standardized tests and 100% on tests targeting the Pittsburgh Urban Mathematics Project (PUMP) curriculum. This curriculum focuses on mathematical analysis of real world situations and the use of computational tools (Koedinger, K and Anderson, J).

In the military domain they have been shown to work as well. The Sherlock maintenance tutor, tested in the late '80s by the Air Force Research Laboratory, trained technicians to troubleshoot faults that were not solved by documented procedures. Thirty-two hours of instruction at technicians' work sites resulted in four years' worth of performance gains, proving the value of individualized simulation-based training (Pokorny, 2012). ONR developed an ITS

in the late 1990s to help sailors who read below Navy minimum standards. Tests conducted with sailors in basic training, and with high school students at risk of dropping out, revealed that students with a history of reading difficulty improved their ability by two grade levels in 40 hours of ITS instruction (Pokorny, 2012). There are a plethora of other examples that show increased learning with the use of an ITS. This does not mean that we need to develop more ITSs and get rid of teachers. ITSs can simply be a tool that teachers use to help instruct or they can be a stand-alone system.

THIS PAGE INTENTIONALLY LEFT BLANK

V. METHODOLOGY

As stated in the introduction, my motivation was to investigate whether I could prove that using an ITS improved training, especially in a surface warfare application. My original idea was to this by creating an ITS for a shiphandling trainer. However, I encountered difficulties getting access to that software, so I decided to investigate an ITS assisting students in conducting a DTE sequence. After surveying software suitable for this task, I decided to use the TAO Sandbox, which was kindly made available for my use by Dr. Allen Munro of the University of Southern California's Rossier School of Education.

A. DESIGN OF EXPERIMENT

After some discussion with my thesis advisors and Dr. Munro we determined that the factor that would be utilized in the experiment would be the ITS itself. That is that there would be two groups of participants in the study: control and experiment. The control group would be given the same three scenarios that the experiment group received; however, in the second of these scenarios, the experiment group would receive treatment with the ITS and the control group would not.

Following more discussion with my advisors and Dr. Munro, I determined that I would use several measures of performance (MOP) that are already present in the TAO Sandbox. These MOPs would include revisiting tracks in a CIEA, evaluating non-ComAir units, directing queries, directing warnings, reporting results of queries and warnings, and reporting intentions. These MOPs will be discussed in depth in Chapter VII.

Defining the inference space is a bit more complicated. According to *Design of Experiments: A Realistic Approach*, "after the experimenter has defined his problem he must make decisions on the limits of the inferences to be made from the results." (Anderson & McLean, 1974, p. 84) I would like to generalize and say that you may infer that the results of my ITS experiment can

be applied to all; however, I cannot do this because ITSs are very complex and often domain specific. Given this, it may only be inferred that the results of the experiment apply to the ITS I designed and only to this ITS. Despite this, successful results from this work would indicate that ITSs are worth further exploration and in need of more designs and experimentation.

In gathering a random selection of the experimental units, I utilized the surface warfare email distribution list as well as sent an email invite to my fellow Modeling, Virtual Environments and Simulation (MOVES) classmates to participate in the study. By opening up to such a large group and utilizing those that responded I was assured a random group of students with varied backgrounds. This would prevent bias towards one specific group, since the surface warfare officers that responded were of varied experience in an air warfare domain as well as the MOVES students, save one, had zero experience in air defense.

Once I received the requisite volunteers, I conducted treatment on the participants based on their availability. The first volunteer was part of the control group. The second volunteer was part of the experiment group. I continued this rotation throughout the volunteers. This allowed the law of chance to enter into the assignment of treatment to all volunteers based on their availability. All volunteers had an equal chance to participate at any given time and no preference was given to any particular volunteer to receive the time slot they requested.

It was difficult to conduct the analysis corresponding to the design before the data was taken given the lack of similar work upon which to base expectations. This is the point where the statistician must write down the mathematical model that has evolved as result of the committee activity in the preceding sections (Anderson & McLean, 1974). There were several items that were discussed at this juncture. One item we discussed was that we needed to block for natural learning that may occur during the experiment. Natural learning, otherwise known as unschooling, is an educational method and philosophy that

rejects compulsory school as a primary means for learning (Wikipedia, 2014). That is, the participants could potentially teach themselves how to conduct an air warfare DTE by simply interacting with the simulator. Utilizing a control group, which did not experience the tutor, provided this blocking. We also discussed the need to determine who would be subject to the experiment. As discussed previously we chose surface warfare officers as they were the immediate target audience of the ITS but we also chose MOVES students in their final quarter of the curriculum, as they were educated in the finer points of simulations and could provide a valuable contribution regarding the use of the simulator and how it was presenting material to the student. Additionally we discussed the size of the experiment group. Due to time constraints we chose to utilize a small sample size which could be run through the repeated measures design experiment. The small sample sizes would be sufficient in demonstrating whether learning was occurring. If it was determined that learning occurred, future work could lend to increasing the sample sizes and take this experiment one step further.

Primarily the simulator itself handled collection of the data. It records values into a text file; however, it aggregates all scores into one file. Therefore, it was necessary for me to copy these text files after each run so that the scores could be determined as to what scenario they originated from. I then averaged the scores from each MOP, from each different scenario and placed them in an Excel spreadsheet for use in the analysis of the data.

Analysis of the data and conclusions will be discussed specifically in Chapters VII and X respectfully. Implementation will not be discussed as this experiment was simply to develop an ITS and conduct a TEE that could potentially be utilized for future work and implementation. As I see it, this ITS could be further built upon, validated and verified, and then pushed to the fleet for future training of war fighters.

B. SCOPE

In understanding the experiment better, it is important to understand the scope of the experiment itself. The scope of this experiment was limited to three air defense scenarios. All of which had the same basic premise behind them. The only thing that changed throughout the scenarios was the origination of the hostile target to prevent the user from gaming the system and expecting the enemy to come from the same spot every time. Furthermore, it was limited to one hostile aircraft, which was placed among four commercial aircraft. The scenario was designed to drive the participant to conduct a DTE culminating in engagement of the target. The weapons posture was set at weapons posture one. That is when mechanical/electrical firing devices are installed in the system; ordnance is loaded; weapons direction system/fire control system may be in auto track/auto engage at command discretion; minimal watchstander actions are required to launch/fire weapons; appropriate for exercise firings or when directed by operational commanders (Federation of American Scientists). Furthermore, the weapons control status was presented as weapons tight. Weapons tight means that weapons systems may be fired only at targets positively identified as hostile. Additionally, each user was given the same mission briefing that there were early indications and warnings that a ship transiting the geographical location of the scenario would be attacked by a country red aircraft.

C. DEVELOPMENT OF ITS

In Chapter IV, I discussed ITSs in depth. I developed an expert model that would watch actions conducted by the user and provide instruction where warranted. This expert model was developed, in part, by utilizing a performance-oriented authoring tool that exists within TAO Sandbox. This tool is called RIDES (Rapid ITS Development Environment). RIDES is used for the construction of tutors that teach students how to operate devices through simulations. This tool is designed to allow a rich learning environment in which students can learn skills by practicing them and receiving feedback (Nkambou, Bourdeau, & Psyche,

2010). For this particular thesis, my intent was to focus on the expert model and develop it for an air warfare DTE sequence. As such I focused on the actions that would be conducted during this sequence by a TAO, generically, as well as the MOPs previously discussed. I will expound on the DTE sequence further in the next chapter, which discusses the methods employed to develop the scenarios. For now, understand that the DTE sequence is a procedure in which a TAO may run through as the attempt to understand the intentions, identification, and threat of a contact.

The actions I felt were paramount to the MOPs as well as the DTE were detecting a threat, attempting to identify the threat, classifying targets as contacts of interests or critical contacts of interest, issuing queries and warnings, reporting results of those queries, and reporting intentions of engagement intentions. Furthermore, it is important to note that if a user does not conduct an action it is generally due to one of two reasons: The first being that they do not know that they are to conduct these actions and the second being that they do not know how to conduct the actions. Based off of these assumptions, I proceeded forward with coding the expert system.

To detect the threat it was paramount that the users turn on their air radars. While it is not usual for a ship to transit in this condition, I felt it relevant to start the user in a condition that their radars were off. It is important as a TAO to always question the information being provided to you by your sensors. From personal experience, I know sensors are not perfect and neither is the output generated by these sensors. False contacts, atmospheric conditions, increased sea state can all attribute to radar screens that are providing unreliable data. By starting the user with their radars off, knowing that there was an air threat and not seeing any air contacts should have made them question what they were seeing. As such the ITS was set on a timer for one minute and if it was determined that the radars were not turned on within this time frame, the ITS suggested to the user that they do so as well as told them how to do so.

Attempting to identify the threat can be done by conducting queries and warnings but also by evaluating the targets air speed, altitude, and course. To ensure that the users were evaluating contacts with all means at their disposal, the expert system was watching to ensure that the users were hooking numerous contacts. If the user did not hook any contacts, the ITS would suggest the users hook a contact as well as suggest they evaluate all contacts on the screen. The ITS would also instruct them on how to hook a contact.

Classifying an aircraft as a contact of interest or a critical contact of interest is not intuitive. In observing the users during the experiment, this assumption was validated. I observed that the users were typically interested in the hostile aircraft; however, they did not always classify targets as such. However, the ITS was coded to suggest that the user classify the contacts as COI/CCOIs when the ITS suggested that they need to query/warn an aircraft. This was a secondary instruction and often overlooked. In hindsight, this should have been done as its own instructional method and not as a subset of queries and warnings. This was done so as there is no set time to classify a target as a COI/CCOI. However, I assumed that when querying or warning an aircraft it was as good a time as any to suggest classifying the aircraft.

Issuing queries and warnings are done at a minimum of once, if time permits. Sometimes pop-up contacts do occur which preclude issuing these warnings; however, I designed the scenarios so that this was not the case. Thus, an expert would conduct a query as the aircraft entered the query zone and would conduct a warning as the aircraft entered the warning zone. The expert system would monitor the student to see if this occurred and after giving the student time to make the proper actions, would prompt them if they had not yet done so. If they did conduct a query or warning, then positive feedback was provided to the user.

Reporting the results of the queries and warnings was did not determine whether the user knew specifically whom to contact, but that the user knew they needed to contact some higher authority. When a query or warning was

conducted, the ITS would suggest to the user that they should report results off the ship. Reporting immediately afterwards is what an expert would do if communications were available, and it was assumed that they were in the simulation. While subjective, most users like this feature, as it was one of the main things that they had forgotten in their short time away from the fleet or field. Always report and do not be the highest ranking member with a secret.

Reporting intentions was the next ability I developed in the ITS. When developing the code, I made several assumptions about the users in this case. These assumptions were that the user would shoot missiles at the target. This assumption was mostly right as 19 of the 20 participants engaged the target. The next assumption was that they would not report their intentions. As such, the expert system monitors when a missile is employed. Following missile employment, the ITS suggests that the user report their intentions off the ship. This allows for command by negation to truly work, if one wanted to incorporate this into the scenario. Command by negation is based around the idea of allowing commanding officers to show their own initiative in executing actions, freeing up their superiors to focus on the bigger picture. Only when the superiors disagree with the proposed actions will they order units not to carry out the actions. This allows ships to perform mostly autonomously but it increases the reporting requirements. Once the user reported their intentions, the ITS would provide them with positive feedback; however, it would also suggest that they report their intentions prior to engagement.

I realize that this expert model is not inclusive of all the actions that users can conduct during an air warfare scenario. There are many different potential user actions which would change the course of events, but creating an ITS which could correctly interpret and remediate for these is outside the scope of this thesis.

The expert model of the ITS was useless without being able to provide the user with instruction and feedback. While my original proposal was to develop only the expert model, it was important that I develop the user interface module

as well. As discussed in Chapter IV, this is the communicating component of the ITS, which controls interaction between the student and the system. This was done by passing messages back and forth between the TAO Sandbox and the ITS. This communication is completely transparent to the user, but it allows modifying the ITS later to respond to different user actions. For the messages from the simulator to the ITS, I had to manipulate the code of the simulator to send messages to the ITS via a TCP/IP connection established by the ITS. These changes are located in Appendix B. For the messages from the ITS to the simulator, a pre-coded format that the simulator understood was required. This required that the output from the simulator be formulated as follows:

chat message this|is|an|example|message|from|the|ITS|to|the|simulator

The simulator would replace the | symbols with spaces as it translated it to the user (USC Center for Cognitive Technology, 2013). This communications protocol developed within the simulator has many other uses as it can be utilized to change the behavior of a contact, deploy contacts, launch torpedoes or missiles as well as many other functions. The input from the tutor to the simulator would be in a different format. One example would be to deploy a hostile aircraft. The location must be specified; however other parameters are optional. This format would be:

Deploy air location<x><y><ID><type><bearing><speed><who-to-attack>

This could be useful for creating a more robust ITS that can react to the student model and provide a more challenging environment for a more experienced user.

VI. TAO SANDBOX SCENARIO DESCRIPTION

I developed three scenarios for this thesis, depicted by figures 4, 5, and 6.. I provided the student with real-world background information for each scenario. Based off of instructional design theory, I did not want to overwhelm or underwhelm the students, but at the same time not make the scenarios so simple as to be trivial. In each, there were several surface contacts and air contacts with which the students could interact. All contacts remained the same throughout; however, the unknown contact (the ultimate contact of interest) origination point changed in each scenario to prevent the student from gaming the scenarios and conducting the same evaluation throughout. The scenario(s) aim was to walk the student through a limited air warfare detect to engage sequence, culminating with a kill.

A. DETECT TO ENGAGE SEQUENCE

The detect to engage (DTE) sequence has many acronyms associated with it. I used a common description, taught at the United States Naval Academy. This acronym is DETITWEE (Detect, Entry, Tracking, Identification, Threat Evaluation, Weapons Pairing, Engagement, and Engagement Assessment). In order to keep this thesis unclassified, I will generalize much of what happens during the DTE sequence.

Detection can happen many ways on a ship, through any number of combat systems that the ship employs or through visual methods. The combat system can be operated by any number of watchstanders. As this thesis is built around an air engagement scenario in TAO Sandbox, detecting an object will occur by the subject, acting as the TAO utilizing a generic radar system. In order for detection to occur the acting TAO must turn the radar systems on. This can be done by clicking on the SWOS emblem at the bottom right corner and selecting the option to turn on all radars. All radars include both surface and air radars in this simulation. While it inaccurate to assume that a ship would transit

without radar on already, having the radars off to begin with was a designed function to help instruct threat evaluation. A TAO cannot simply trust what they are seeing; they must always question what they are seeing on their radar scope. Being briefed that this is an air warfare scenario and not seeing any aircraft, should prompt the student to question why this is the case. There is also an option to turn on all sonars, in the event that this thesis were to be translated to an Anti-Submarine Warfare (ASW) or Mine Warfare (MW) scenario. While there are instances that radar and sonar may be turned off, we will assume that the scenarios presented to the user will not require the ship to turn off any of these.

Entry is the process of inputting the data into the Naval Tactical Data System (NTDS) that Navy ships utilize. For many ships, this can be as simple as “hooking,” or selecting, the target. By hooking the target the user assigns it a Target ID number. In TAO Sandbox (version 3.90) hooking is done by clicking the ‘A’ next to the target; which will provide a pop-up menu. From the menu the user will select ‘Hook’. Once this happens, a kinematics box pops up and there will be a number displayed below the target and that correlates to that box. This kinematics box will contain the track number, which correlates to the number below the target, as well as range, bearing, speed, course, and altitude. A limitation of the simulator provides that only one contact can be hooked at a time. This is inaccurate as there may be multiple contacts that are hooked by a TAO.

Tracking is the process of accurately determining the target’s position. As with many of the newest combat systems you will only have to hook the target and the computer will do this for you. In older systems, a plot may have been used. In order to maintain a plot, a target would be assigned a number and would be tracked by taking several hits on the target over a period of time.

Identification is the process of identifying a potential target. This can be done on the ship by utilizing Identification Friend or Foe (IFF), which can be queried using special equipment onboard. According to the Navy Department Library, “there are two basic components in any identification system at present in use: the transponder, carried by the aircraft or ship to be identified and the

interrogator-responder, or questioning device, located aboard a ship, at a ground station, or in another aircraft.” (The Navy Department Library) This is essentially a handshake between units. If the transponder sends back a pre-designated code, the aircraft being queried by the IFF system can be identified as friendly or as a commercial aircraft. This is referred to as checking modes and codes. There are other ways to identify a contact, such as using a helicopter or other known friendly unit for a visual check, commonly known as a visual ID. In the case of a commercial airliner they typically have airways that they fly in that are known to the ship, which can be used along with their altitude, speed, and transponder code to classify them as a commercial airliner. This does not alleviate the TAO of continually checking potential threats, as there are ways to spoof codes. The designated airways are also public knowledge so it is possible that a potential threat could fly in those airways, at a similar speed and altitude of a commercial jet in an attempt to gain proximity to the ship before turning inbound and descending rapidly, very similar to a dive bomber like approach.

Threat Evaluation is the process of determining the relative degree of threat (threat priority); based on position, approach, ID, range of weapons, and time remaining to effectively engage. This is usually the most complicated phase of the DTE sequence. There are many considerations that come into play for a TAO when evaluating a potential threat. For example, the TAO might be interested in a target if there was 000 relative bearing, with a decreasing range also known by many SWOs as CBDR (Constant Bearing, Decreasing Range). They might also be interested in the target if it entered their query or warning zones. The query zone is a pre-defined zone in which an aircraft must be queried as to their identity and intention. The warning zone is a pre-defined zone in which an aircraft must be warned that in the event they do not identify themselves or their intentions, and they continue to approach the vessel that the ship may use deadly force in order to protect itself. This is essentially a red line, after which the TAO may engage the aircraft. In TAO Sandbox, the ranges may be displayed by clicking on the ‘A’ next to ownship (the TAOs vessel), which will generate a pop-

up menu, much like hooking did. The TAO must select 'Air Defense Tasks' and then 'Show Query Range' or 'Show Warning Range'. Once this occurs the zones will be displayed on the TAO Sandbox screen. The final zone, which can also be displayed in a like manner of the query and warning zones, is the vital area. The vital area is the zone in which deadly force will actually be used if a target is deemed a threat. These ranges are typically located in the OPTASK Air Defense, which is a classified document. For TAO Sandbox, these ranges have been generalized to keep the simulator unclassified and are easily manipulated to demonstrate actual ranges.

Weapons Pairing is the process of assigning optimum weapon for use against a given threat based on threat priority and available assets. For this thesis the TAO will be stationed inboard a DDG Flight 2. According to the Navy, a DDG Flight 2 has Standard Missiles (SM-2MR); Vertical Launch ASROC (VLA) missiles; Tomahawk; six MK-46 torpedoes (from two triple tube mounts); Close In Weapon System (CIWS), five MK 45 Gun, and the Evolved Sea Sparrow Missile (ESSM) (United States Navy, 2013). The numbers may vary by mission and weapons load out but for this scenarios presented to the user in the TAO Sandbox they have thirty Harpoons, two SH-60B helicopters, as well as chaff, nixie, and surface to air missiles (SAM). For the particular scenarios, they were designed with air warfare in mind so the proper weapons pairing would be the SAM.

Engagement is the process of employing the ships weapons systems. In the scenarios, the TAO should employ the surface to air missile (SAM). To do so, the user would select their 'A', then selecting 'Launch or Drop', and then selecting 'SAM'.

Engagement Assessment, also referred to as a battle damage assessment, is the process of monitoring weapon systems return information to determine whether the engagement was a success or if reengagement is necessary. The probability of kill for these particular scenarios is 1, meaning that every missile will hit and kill its intended target. This is depicted by a red asterisk

on the display, where the target had been prior to engagement. To successfully complete an engagement assessment in these scenarios it was required that the user send a message off ship either communicating a situation report (SITREP) or a battle damage assessment (BDA) to higher authority.



Figure 4. Scenario 1

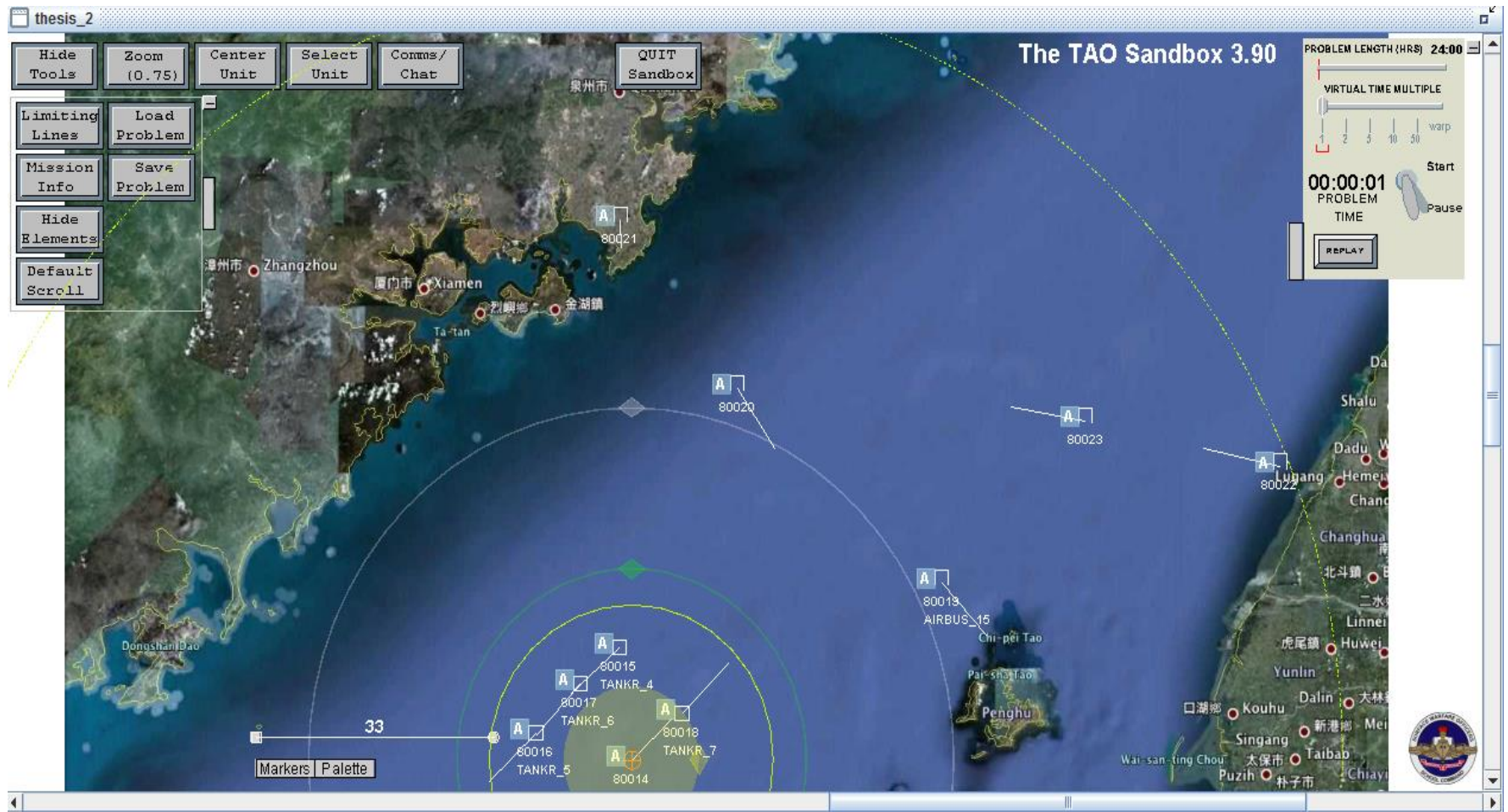


Figure 5. Scenario 2

VII. EXPERIMENT DATA AND ANALYTICAL RESULTS

Collected data is useless if you do not analyze it to tell a story about what is gained from the experiment itself. Using statistics, I will summarize the data, analyze the data, and draw conclusions based on the evidence provided by the data. My analyzing the data I can support, or refute, my hypothesis as well as to determine if the results are significantly significant. Significantly significant results means that it is unlikely that my results are due to chance (Cherry, 2014).

A. SUMMARY OF DATA

Appendix F contains the data collected from the output of TAO Sandbox. TAO Sandbox scores certain actions, which it calls NTAs. These are measures of performance that are output to a text file by the simulator. Below are the breakdowns of the measures of performance:

NTA 1: Revisiting Tracks in Classification, Identification, and Engagement Area (CIEA). Good scores are granted for accessing the kinematics of units within the CIEA. Poor scores are given for each hostile or unknown that is not hooked within a short time after entering the CIEA.

NTA 2: Evaluating Non Commercial Aircraft (ComAir) Air Units. Good scores are given when approaching unknown or hostile air contacts are labeled as a contact of interest (COI) or a critical contact of interest (CCOI). If an approaching unit gets close enough without being labeled as a COI or CCOI then a bad score is recorded for this measure.

NTA 4: Directing Queries. NTA4 has been implemented by giving good scores to students who query approaching aircraft. Highest scores are given for querying soon after crossing into the query range (45NM for this experiment) but before entering the warning range. Lesser, but still good scores are given for doing queries outside query range, and still lower scores are given for not querying until after the approaching aircraft has entered the warning range.

NTA 5: Directing Warnings. NTA5 scores report performance on issuing warnings to potential air threats. Highest scores are given for warning soon after the potential threat has crossed the warning range. Lesser scores are given for issuing warnings outside the warning range. And still lower scores are given for not warning until after the CIEA has been entered.

NTA 6: Reporting Results of Queries and Warnings. When a query or warning response occurs, the Sandbox notes the time of the response and waits to see whether the results are reported to the air defense controller (ADC). If the report is made quickly, a high score is reported. A slower report results in a lower score. Not reporting at all results in the lowest score.

NTA 7: Reporting Intentions. Students receive positive scores for reporting intentions. If a student carries out a tactical action, including using chaff or firing a SAME, without first reporting the intention, a low score is recorded.

The numerical scores that are recorded are values between zero and one. An excellent score for a measure would be a 0.9 or higher; a very poor score would be 0.1 or lower (Munro, 2013).

B. ANALYSIS OF DATA

To analyze the data, I considered using a two-sample t-test to compare the means. The two-sample t-test is a hypothesis test for answering questions about the mean where the data are collected from two random samples of independent observations, each from an underlying normal distribution (Columbia Center for New Media Teaching and Learning (CCNMTL), 2003).

There are certain assumption and conditions that must be met in order to conduct this test. These are the independence assumption, the normal population assumption, and the independent groups assumption. The independence assumption states that the data in each group “must be drawn independently and at random from a homogeneous population, or generated by a randomized comparative experiment” (De Veaux, Velleman, & Bock, 2008, p. 583). This particular assumption is met by the data being collected from a

randomized comparative experiment. The normal population assumption states that both groups' data must be normal or nearly normal. This matters in this case as the sample size for both groups is small. This assumption can be best be clarified by a histogram. The histogram, below, of the control group for NTA1 demonstrates that the data cannot be normalized. Therefore, I was unable to use the two-sample t-test.

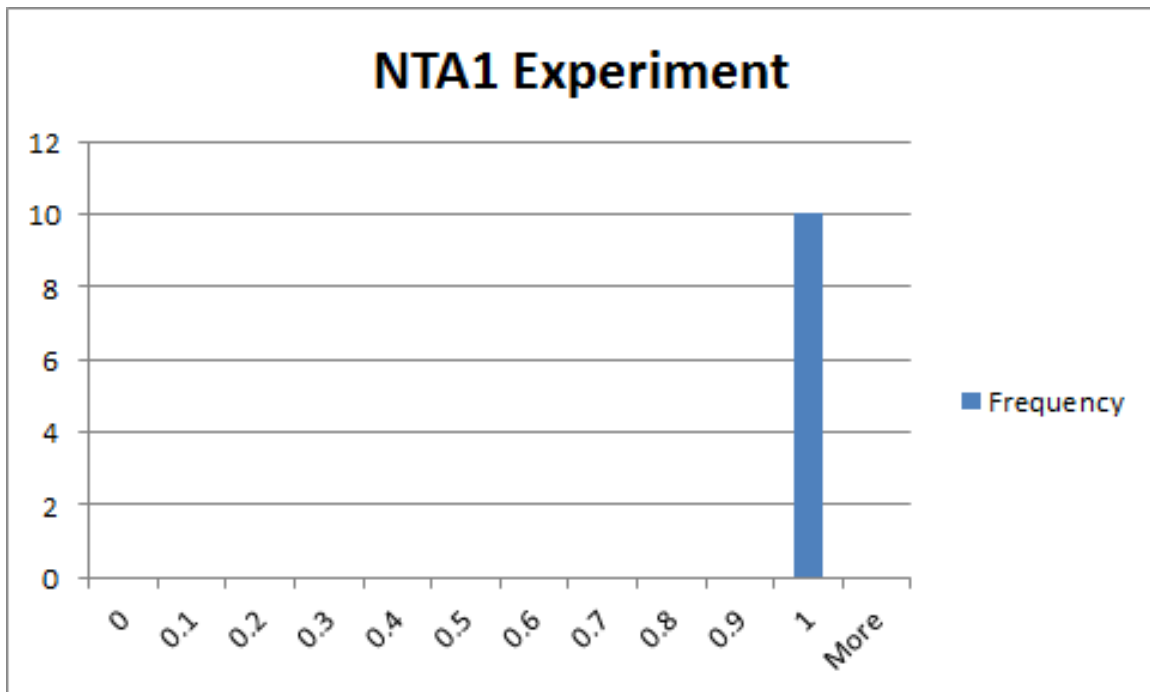


Figure 7. NTA 1 Histogram of Experiment Group

As mentioned above, the data was not shown to be normal. As such, I attempted a Wilcoxon rank sum test, also known as a Mann-Whitney test. The Wilcoxon rank sum test first ranks the combined sample from the two groups together from smallest to largest, assigning each observation a rank from 1 to N (N being 12 in this instance) (De Veaux, Velleman, & Bock, 2008). The test statistic, W, is the sum of the ranks of the first group (control). In this case, the hypothesis test is based on a couple of facts that can be derived mathematically. When the null hypothesis is true, the test statistic W has a mean $\mu_W = \frac{n_1(N+1)}{2}$

and variance $Var(W) = \frac{n_1 n_2 (N+1)}{12}$. For all but small sample sizes (both larger than 7), we can use a z-test to test the hypothesis, where $z = \frac{W - \mu_W}{SD(W)}$ where the $SD(W) = \sqrt{Var(W)}$. My sample size of participants in each group is ten, with six different MOPs. Since I am breaking these down into MOPs, $n_1 = 6$ and $n_2 = 6$. While this makes the sample size less than seven, I will continue on cautiously.

Rank	Data (Avg of all runs)	Group
1	0.173	Control NTA 5
2	0.206	Experiment NTA 5
3	0.209	Control NTA 7
4	0.326	Experiment NTA 7
5	0.34	Control NTA 4
6	0.429	Control NTA 6
7	0.459	Experiment NTA 4
8	0.619	Experiment NTA 6
9.5	0.704	Control NTA 2
9.5	0.704	Experiment NTA 2
11	0.843	Control NTA 1
12	0.937	Experiment NTA 1

Figure 8. Wilcoxon rank sum of averages for all three runs

Using these formulas on the above data we find $W = 35.5$, $\mu_W = 39$, $Var(W) = 39$, $SD(W) = 6.245$, and $z = -.560$. If we consult the Z table, we find that the one-sided P-value is .2877. However, since we are interested in deviations on either side of the null hypothesis the two-sided P-value is .5754. I also chose to use an alpha level of .10. An alpha level, also known as a significance level, is the probability value below which the null hypothesis is rejected. With the Wilcoxon rank sum test the null hypothesis is that there is no difference in the observations between two groups. Given that the two-sided P-value is greater than .1, I am willing to retain the hypothesis that there is no

difference in the observations between the control group and the experiment group. This means there is no statistical difference showing that the ITS does effect some learning upon the users. As these are the averages over three runs, it is possible that the control group started out with higher first runs, which could be skewing the data. A better way to investigate statistical significance would be to check the deltas. That is the differences between runs 1 and 3.

The Wilcoxon rank sum test is a less powerful statistical test than a two-sample t-test, as it does not utilize all information from the data (De Veaux, Velleman, & Bock, 2008). Because of this and since I am testing a different data set I will first attempt a two-sided t-test on the deltas. The assumptions remain the same as above for a two-sample t-test: random, independent, and normal data. Because this data is coming from the same test groups the random and independent assumptions remain. I will check normalization by looking at the deltas from NTA 1.

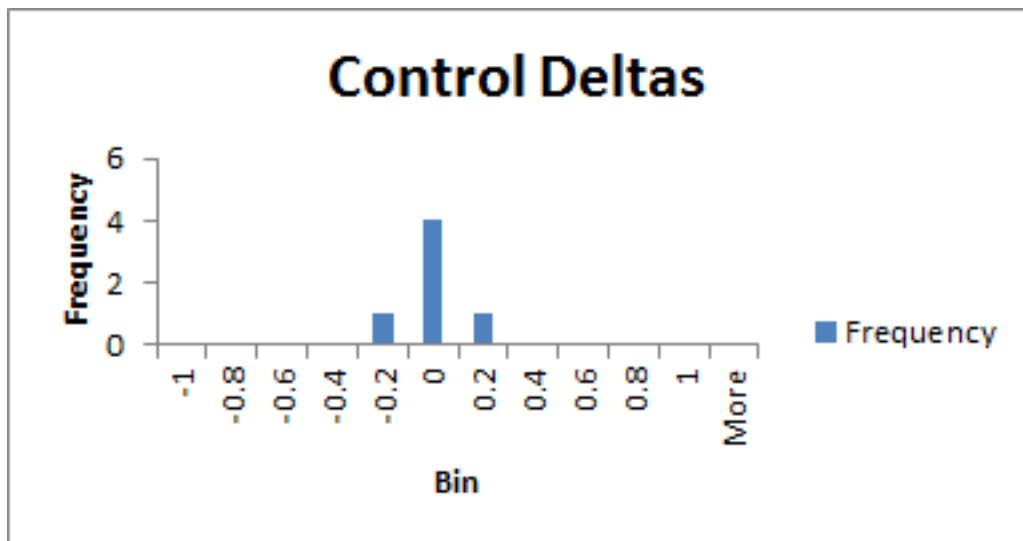


Figure 9. Histogram of Control Deltas

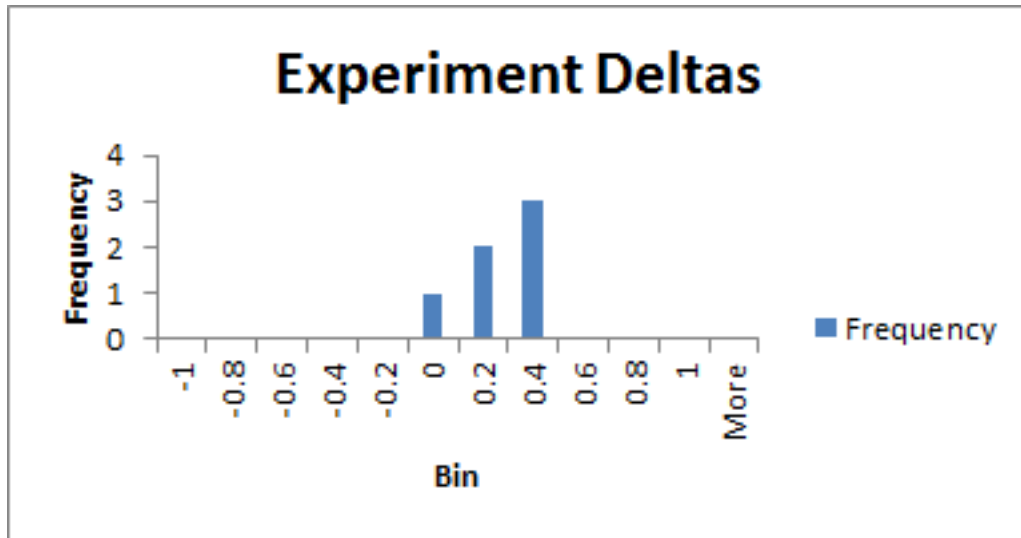


Figure 10. Histogram of Experiment Deltas

Based on the above histograms, it appears that the control group deltas are normal; however, the experiment deltas are not normal. Due to failing the normal assumption required of the two-sample t-test, I will now conduct a Wilcoxon rank sum test. Below is the delta data:

Rank	Data (Delta)	Group
1	-0.249	Control NTA 4
2	-0.1385	Control NTA 5
3	-0.13	Experiment NTA 5
4	-0.088	Control NTA 2
5	-0.027	Control NTA 6
6	-0.006	Control NTA 1
7	0.088	Experiment NTA 2
8	0.14	Experiment NTA 1
9	0.1505	Control NTA 7
10	0.325	Experiment NTA 7
11	0.337	Experiment NTA 4
12	0.351	Experiment NTA 6

Figure 11. Wilcoxon rank sum of deltas between runs one and three

Using the Wilcoxon rank sum formulas on this data we find $W = 27$, $\mu_W = 39$, $Var(W) = 39$, $SD(W) = 6.245$, and $z = -1.92154$. If we consult the Z table, we find that the one-sided P-value is .0274. However, since we are interested in deviations on either side of the null hypothesis the two-sided P-value is .0548. Based off of this P-value and an alpha level of .1, I reject the null hypothesis that there is no observable difference in the deltas of the control and experiment group. Thus, there is statistical evidence that the ITS was effective in effecting teaching on the study participants.

To better understand the analysis, I feel it relevant to include the mean and standard deviations of the deltas of each group in regards to each MOP. The mean is the average of the numbers. To calculate this we will add each delta from each user in their respective groups and divide by 10 (number of participants in each group). From the data presented in the beginning of this chapter, we find the means to be:

DELTA				
MOP	Control Sum	Control Mean	Experiment Sum	Experiment Mean
NTA 1	-0.06	-0.006	1.4	0.14
NTA 2	-0.88	-0.088	0.88	0.088
NTA 4	-2.49	-0.249	3.37	0.337
NTA 5	-1.385	-0.1385	-1.33	-0.133
NTA 6	-0.27	-0.027	3.51	0.351
NTA 7	1.505	0.1505	3.33	0.333

Figure 12. Means of deltas between runs 1 and 3 for control and experiment groups

The standard deviations presented found by utilizing the formula $s = \sqrt{\frac{\sum(y-\bar{y})^2}{(n-1)}}$. Utilizing this equation we get the following:

Standard Deviations		
MOP	Control	Experiment
NTA 1	0.364910948	0.289405521
NTA 2	0.499528667	0.306248991
NTA 4	0.24259935	0.22326392
NTA 5	0.153261759	0.179040964
NTA 6	0.156492811	0.412384529
NTA 7	0.280974989	0.381535057

Figure 13. Standard Deviations of MOPs for control and experiment groups

Additionally I believe it relevant to provide some detail about the breakdown of the groups to demonstrate that they were balanced:

	SWO	SWO %	MOVES	MOVES %
Control	6	60%	4	40%
Experiment	6	60%	4	40%

Figure 14. Breakdown of users in each group

The above data show that there were an equal number of surface warfare officers in each group and an equal number of MOVES students in each group. Surface warfare officers accounted for 60% of all participants and MOVES students accounted for 40% of all participants. It is completely coincidence that the numbers worked out this way, as the participants were randomized based upon their time availability; however, it does demonstrate the balance of the groups.

THIS PAGE INTENTIONALLY LEFT BLANK

VIII. DEMONSTRATION OF CONCEPT

In order to fully understand what the tutor is doing, I felt it important to include a demonstration of concept. This will include snapshots of the tutor in action and discuss what the simulator and tutor are doing at each juncture. I will use scenario 2 for this demonstration, as it is the same scenario that the experiment group utilized the ITS in.

In Figure 15, the ITS is introducing itself to the user. Prior to this occurring a connection had to be made between the ITS and the simulator. This was done utilizing a TCP/IP connection and in order to ensure the tutor is the connecting entity, a username and password is exchanged. In order for this to occur the btl.prp file had to be modified to add the tutor.

In Figure 16, the ITS is informing the user that they should turn on their radars. Once the simulation is started the ITS starts a clock and gives the user one minute to turn on radars before suggesting to do so. The ITS is looking for a message from the simulator that the radar has been turned on. When this action happens in the simulator, it sends a message of RADAR.

In Figure 17, the ITS is informing the user that they can hook the contacts to learn more about each contact by bringing up the contacts kinematics. The user has a minute and a half from starting the tutor to do this. The ITS is once again looking for a message from the simulator that a contact has been hooked. When this action happens in the simulator, it sends a simple message of HOOKED.

In Figure 18, the ITS is informing the user that they should evaluate all contacts on the screen and not focus on one contact. The tutor allows for two minutes from the start time for this to occur. Similar to the hook and radar functionality, the ITS is looking for a message from the simulator that the user is changing contacts, which is done by the simulator sending a message of CHANGED. In any of the RADAR, HOOKED, or CHANGED functionalities, if the

user performs these actions prior to the tutor time limits, the tutor will not send the message. This is because the actions are in agreement with what the expert model is doing.

In Figure 19, the tutor is informing the user that they have an unknown contact in their query zone. The simulator is using a watcher for this functionality. The watcher is looking at the distance of the unknown contact from the user's unit. Once the unknown contact enters the query zone, the watcher is triggered and sends a message to the ITS. The ITS then starts a timer to give the user time to conduct this action on their own, to allow for the user to do this on their own without guiding the user through every step. This is in line with the cognitive theories discussed in Chapter II. Additionally at this point the tutor suggests to the user that they mark the contact as a COI. Once again, this is the exact spot that this action must be done, but this point of instruction allowed for this metric to be addressed.

In Figure 20, the tutor is acknowledging that the user queried the unknown vessel and providing immediate positive feedback reinforcing that this was the correct action. The tutor now utilizes this juncture to remind the user that they should report the result of the query off ship to ADC. The tutor knows that this query has been conducted because the simulator sends a message to the tutor letting it know that the query has been conducted.

In Figure 21, the tutor is providing immediate positive feedback to the user on their reporting the query results to reinforce that this action was the correct one. The tutor is aware that this message was sent due to the simulator notifying the tutor.

In Figure 22, the tutor is suggesting that the user warn the unknown vessel. My first attempt was to have another watcher for this functionality; however, the second watcher did not want to work properly in beta testing. As such, this knowledge is known by the tutor only through a time distance formula ($D = ST$). Utilizing the speed of the aircraft which was hardcoded into the ITS and

the known distance between the query zone and warning zone, the tutor utilized the time that the aircraft would require to travel through the query zone to the warning zone and added another 45 seconds to allow for the user to conduct the action on their own. The simulator sends a message to the tutor when this action has been completed. If the user completes the action prior to the allotted time then the tutor does not suggest any actions to the user. Much like the query functionality, once a warning is issued the tutor provides positive immediate feedback and suggests that the user report the results off the ship.

Figure 23 simply demonstrates the simulator automatically classifying the contact as hostile. This change in symbology along with the knowledge of weapons posture and status should drive the user to shoot. There is currently no functionality that would indicate to the tutor that this classification has been done; however, based on personal observations during the experiment this functionality should be implemented, as a staggering number of participants were not aware of what this weapons posture and status meant.

Figure 24 demonstrates the initiation of shooting. Once the user determined they were going to shoot and conducted the actions to fire a SAM, the simulator sent a message to the tutor letting it know that a missile had been fired.

Figure 25 shows the tutor suggesting the user report their intentions off of the ship at this time. Had the user done this prior to shooting the missile, positive immediate feedback would have been given to let the user know this was the correct procedure. However, in this snapshot, the user had not reported their intentions off ship prior to shooting the SAM. The tutor is allowing them to do that now and suggesting that they do it before firing next time. Further the tutor takes this moment to remind the user to report the results of the engagement off ship.

Figure 26 shows the user providing a SITREP following a successful engagement. The simulator sends a message to the tutor so the tutor is aware

that this has been done and the tutor provide immediate positive feedback to the user to reinforce that this was the correct procedure.

Figure 27 is demonstrating the communication between the tutor and the simulator. The messages sent to the simulator from the tutor are preceded by the words received. These are usually messages to the user from the ITS that are communicated through the simulator. This is representative output of the user interface module that was discussed in Chapter IV. Messages without the preceding word received are messages from the simulator to the tutor to allow the tutor situational awareness as to the actions being conducted to the user. This is representative output of the tutoring module that was discussed in Chapter IV.

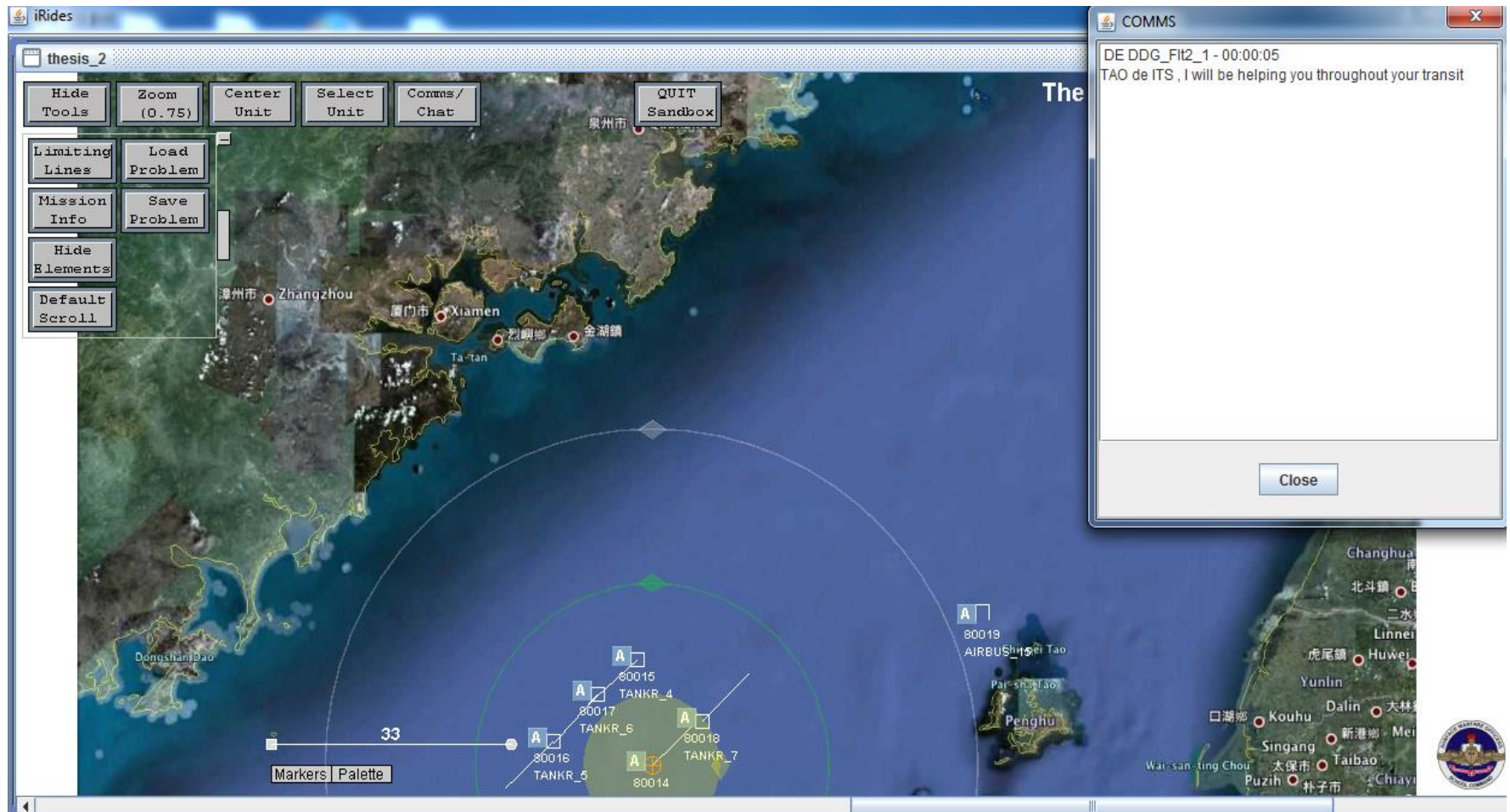


Figure 15. ITS introduction

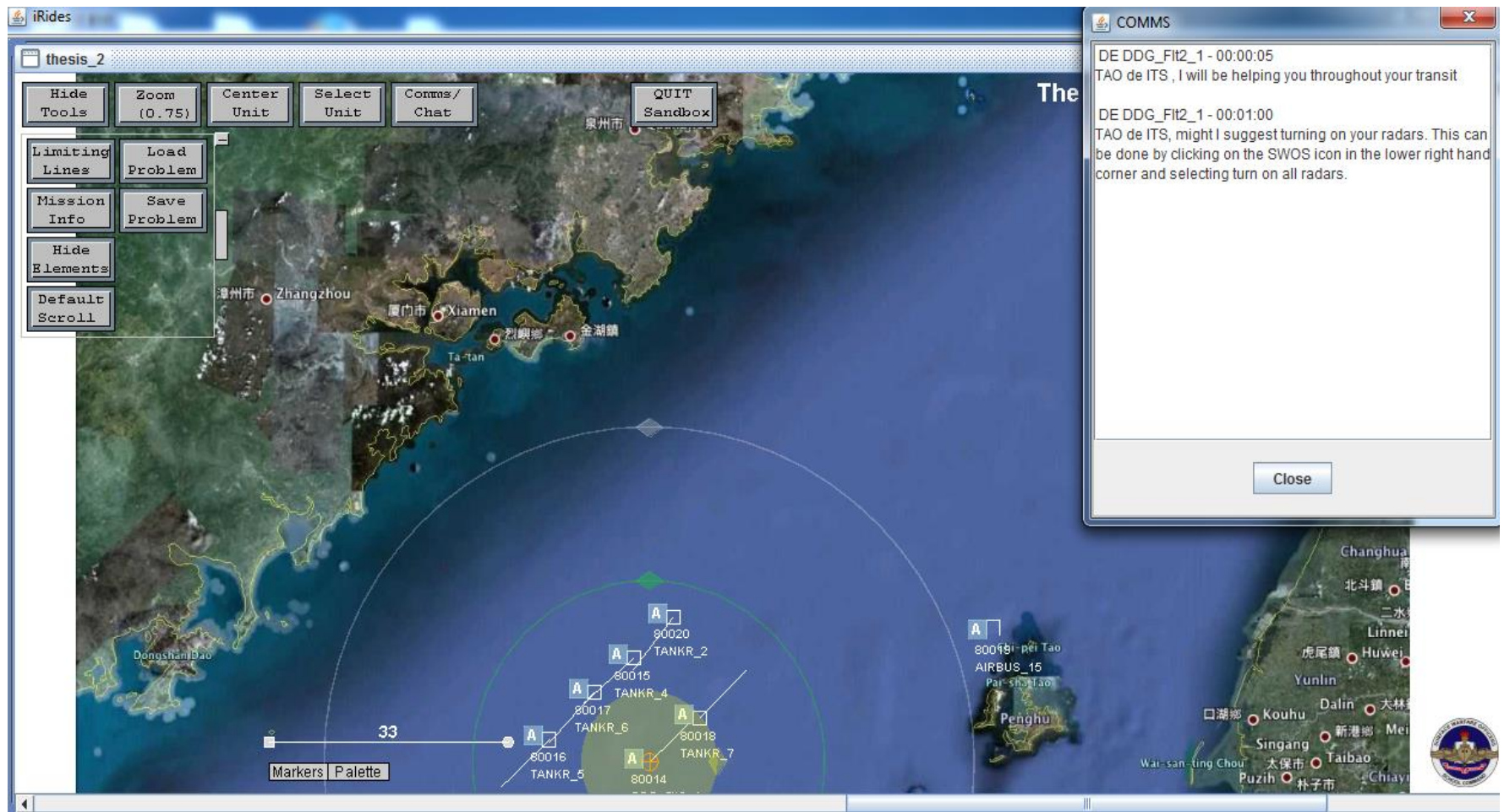


Figure 16. ITS suggesting turning RADAR on

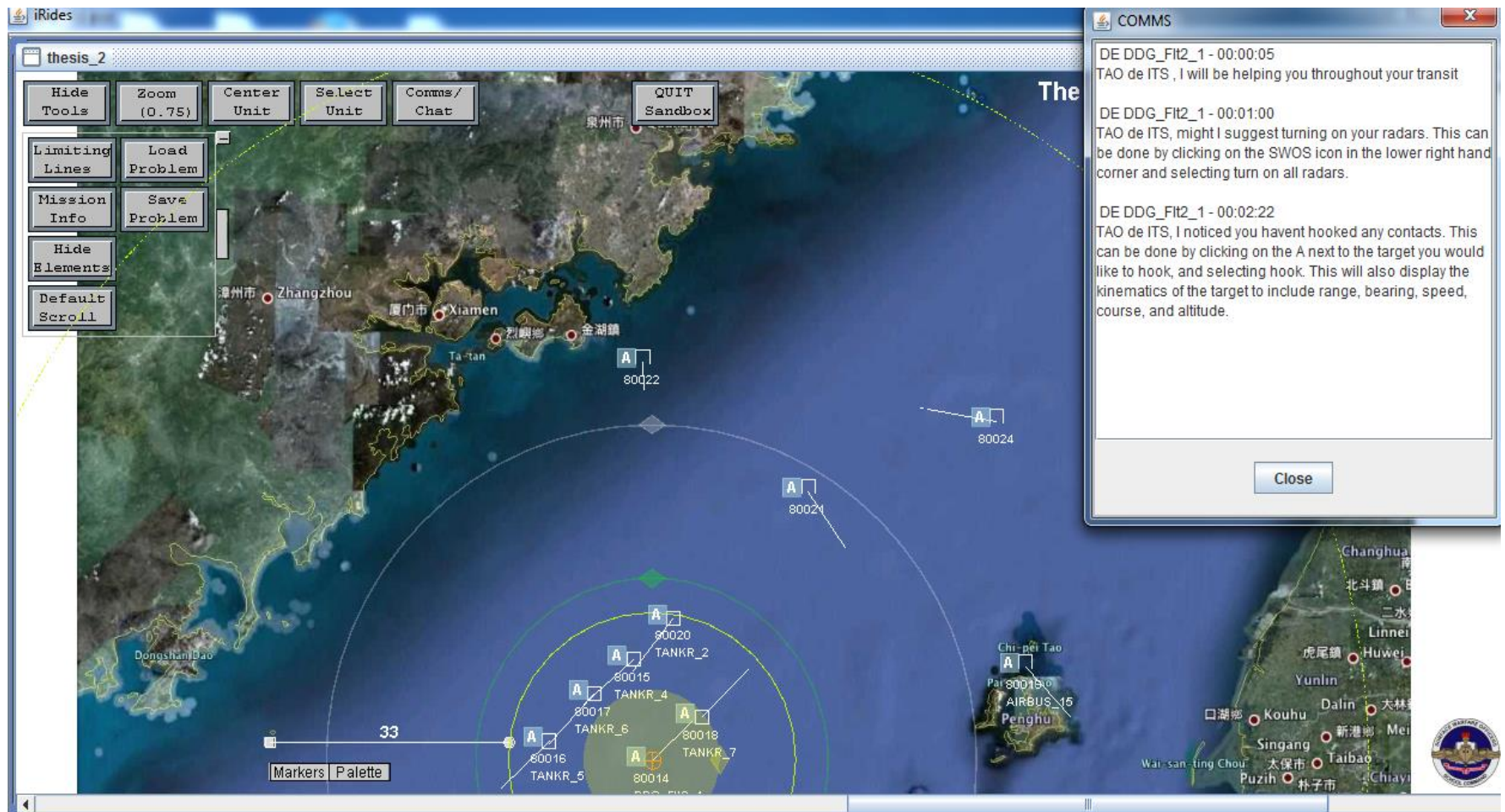


Figure 17. ITS suggesting HOOK action to user

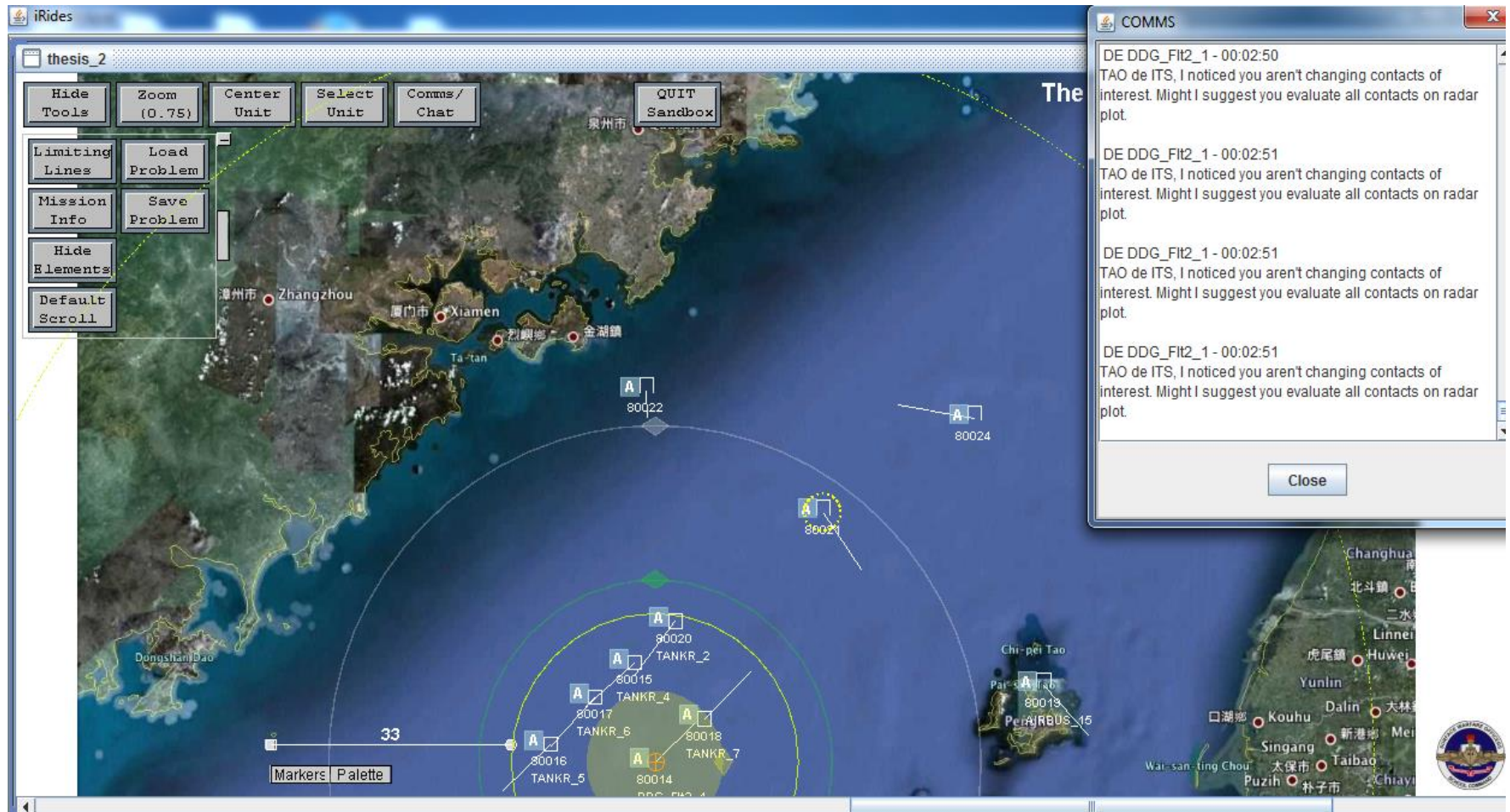


Figure 18. ITS suggesting evaluation of all contacts

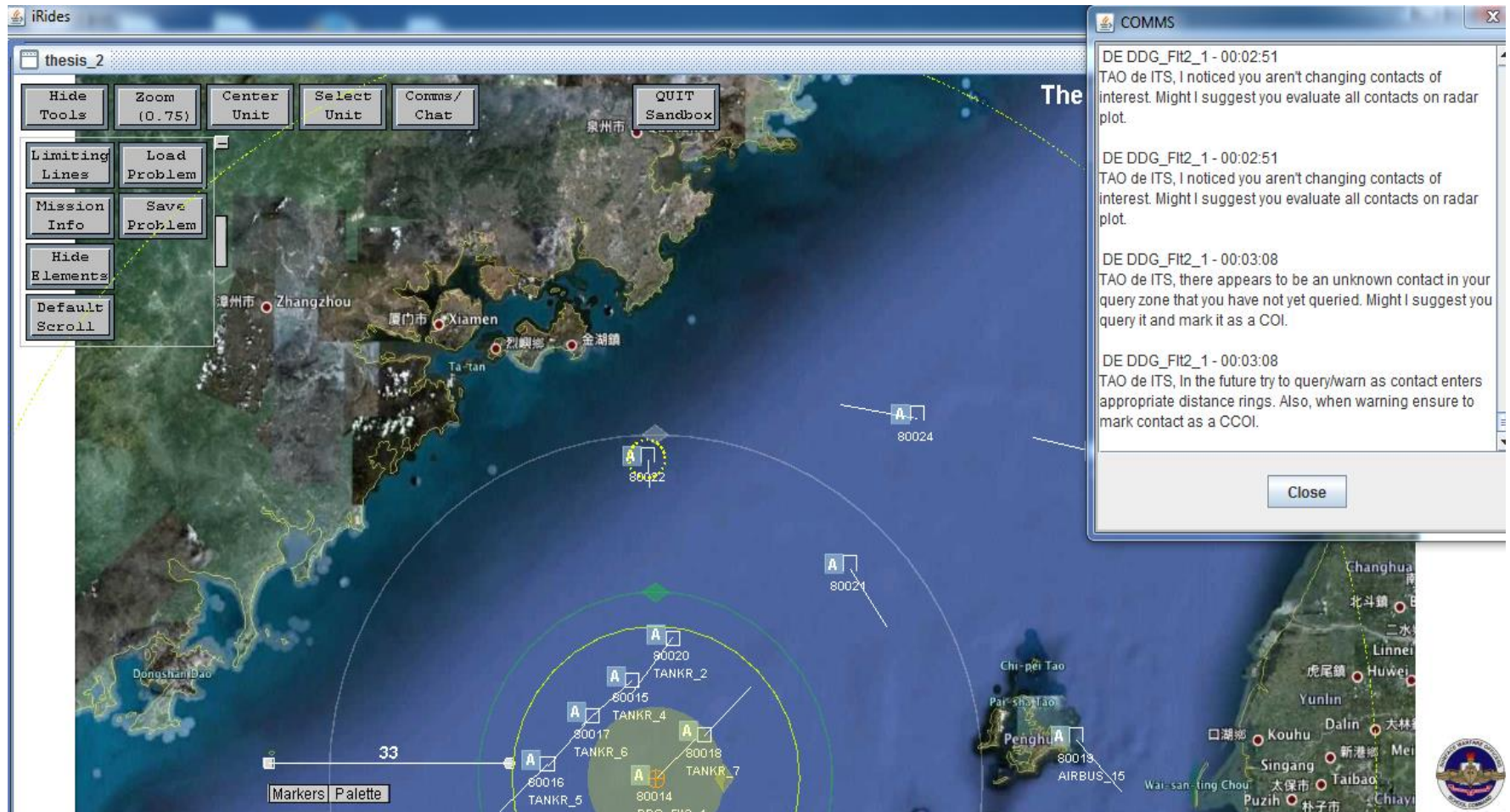


Figure 19. ITS suggesting user query unknown contact

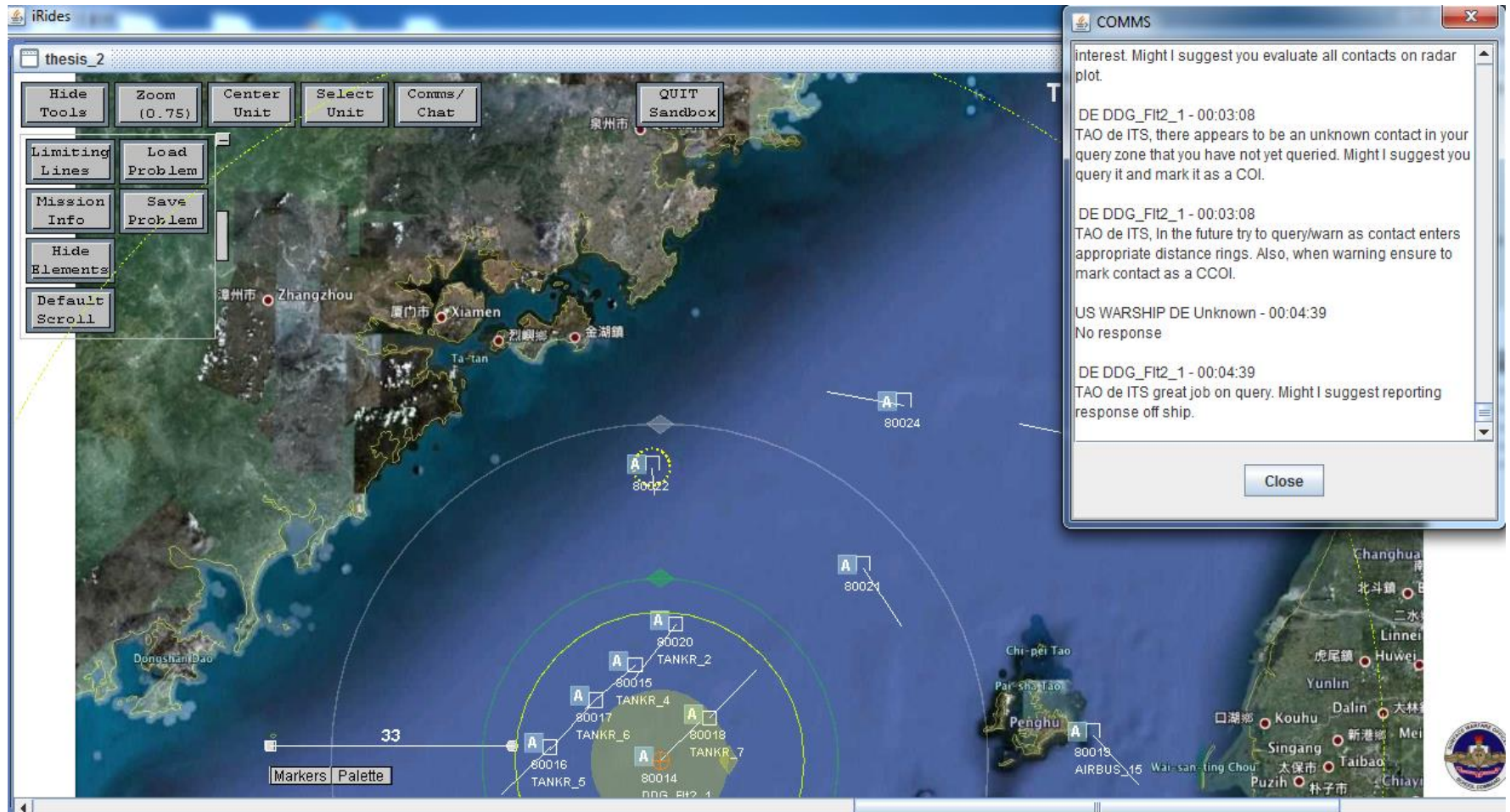


Figure 20. ITS providing feedback to user about successful query

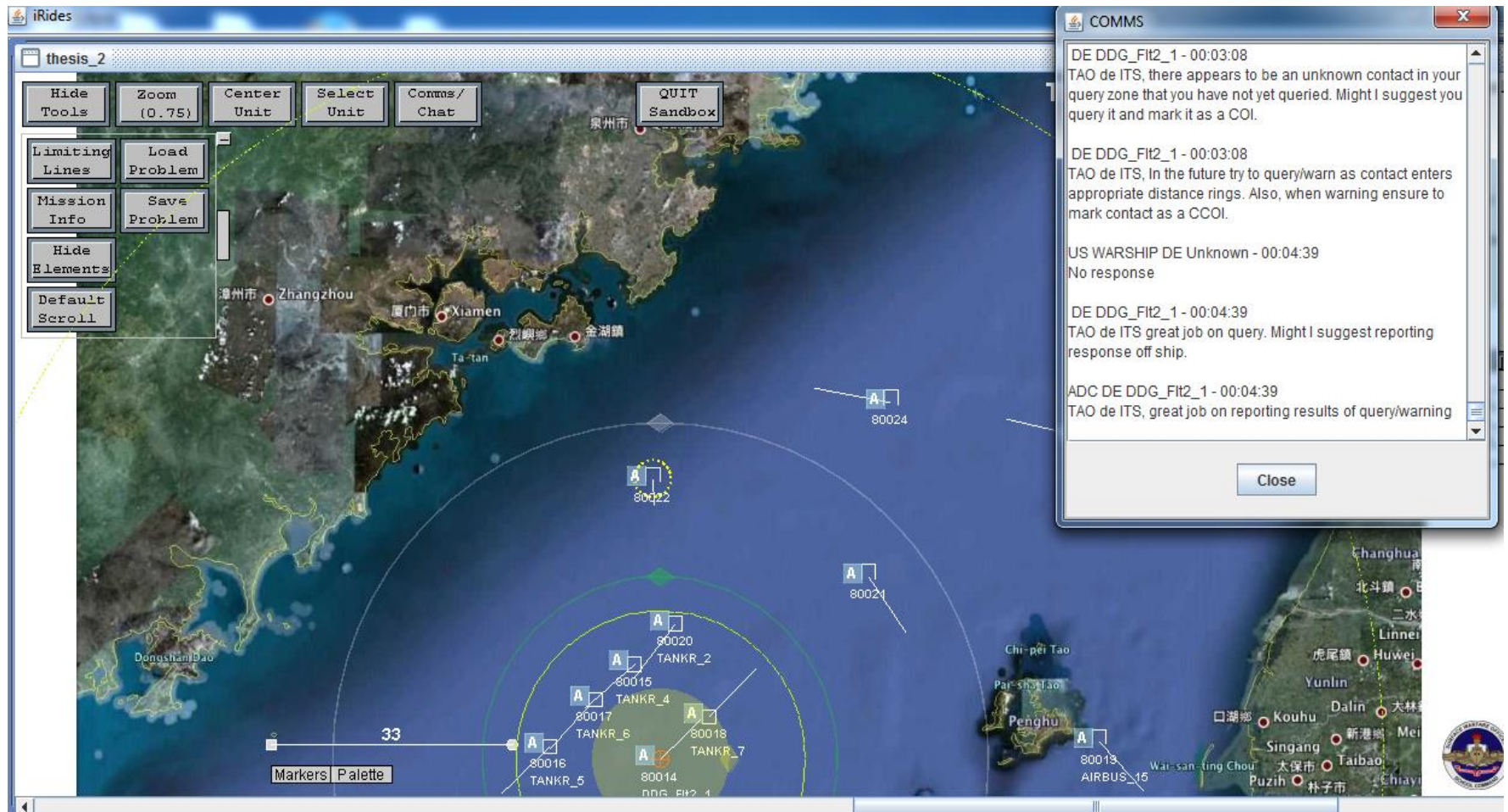


Figure 21. ITS providing feedback on results of reporting query results

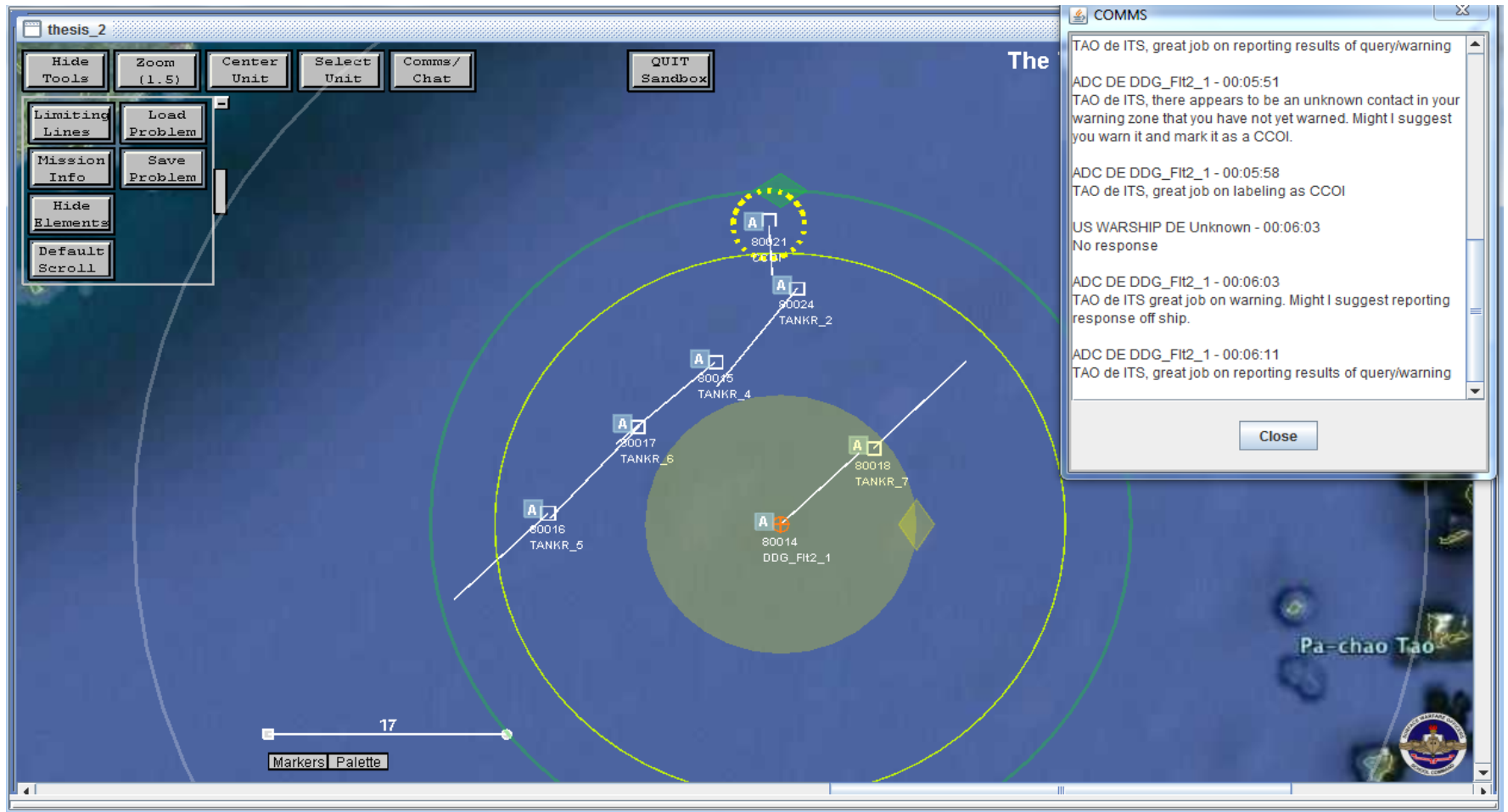


Figure 22. ITS suggesting the user warn unknown contact

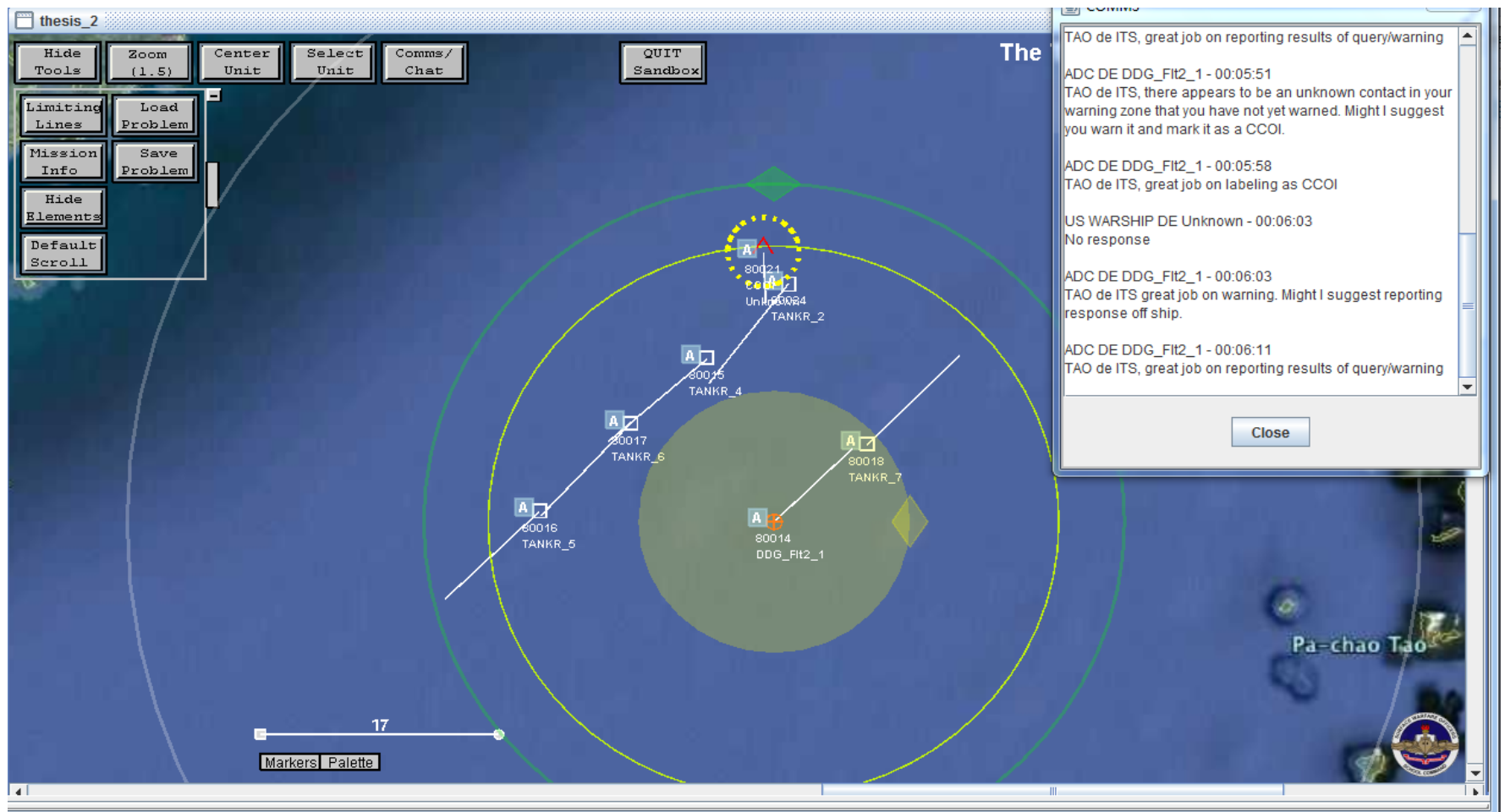


Figure 23. Simulator classifying unknown contact as hostile

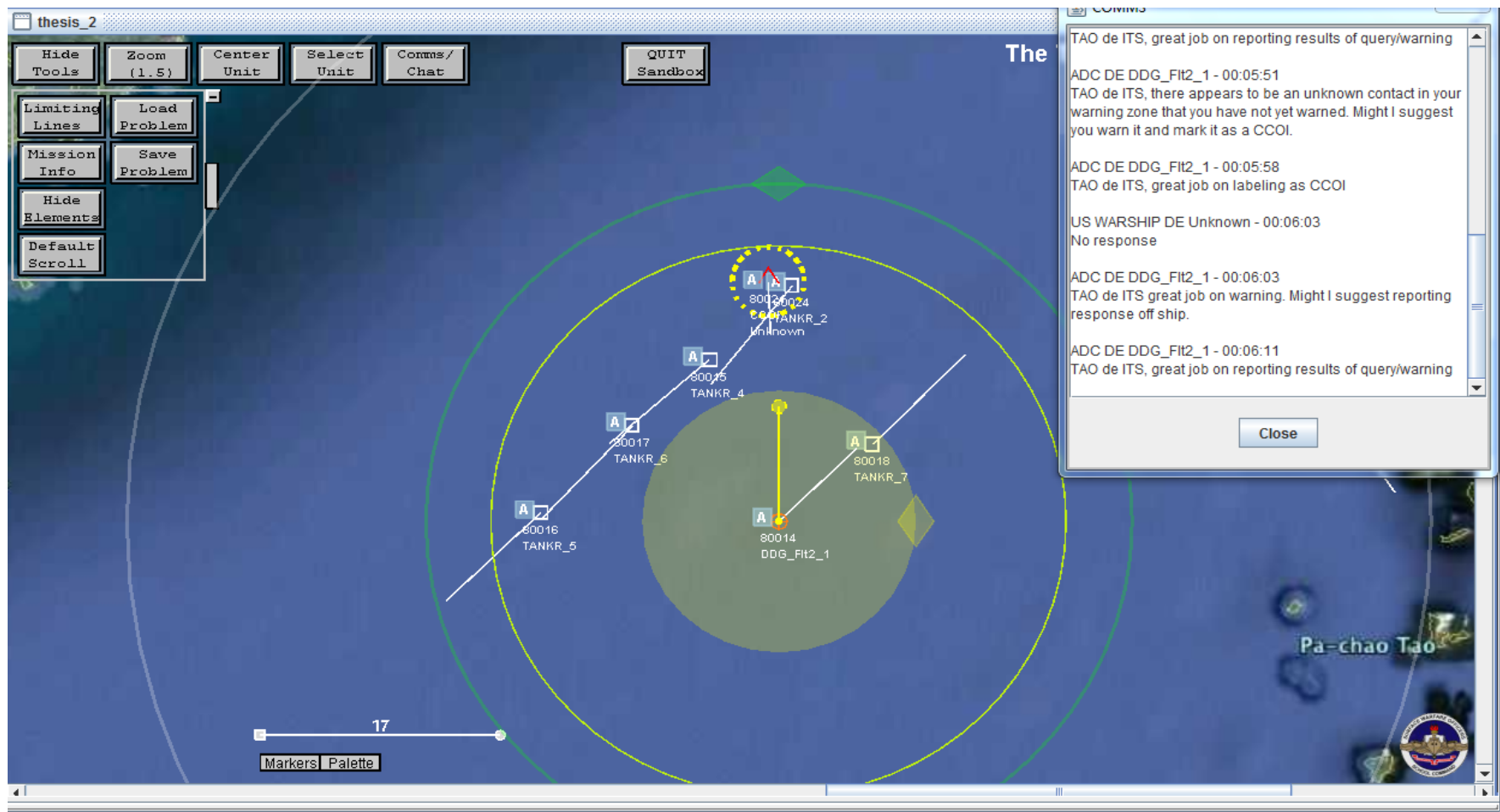


Figure 24. User initializing engagement of unknown contact with SAM

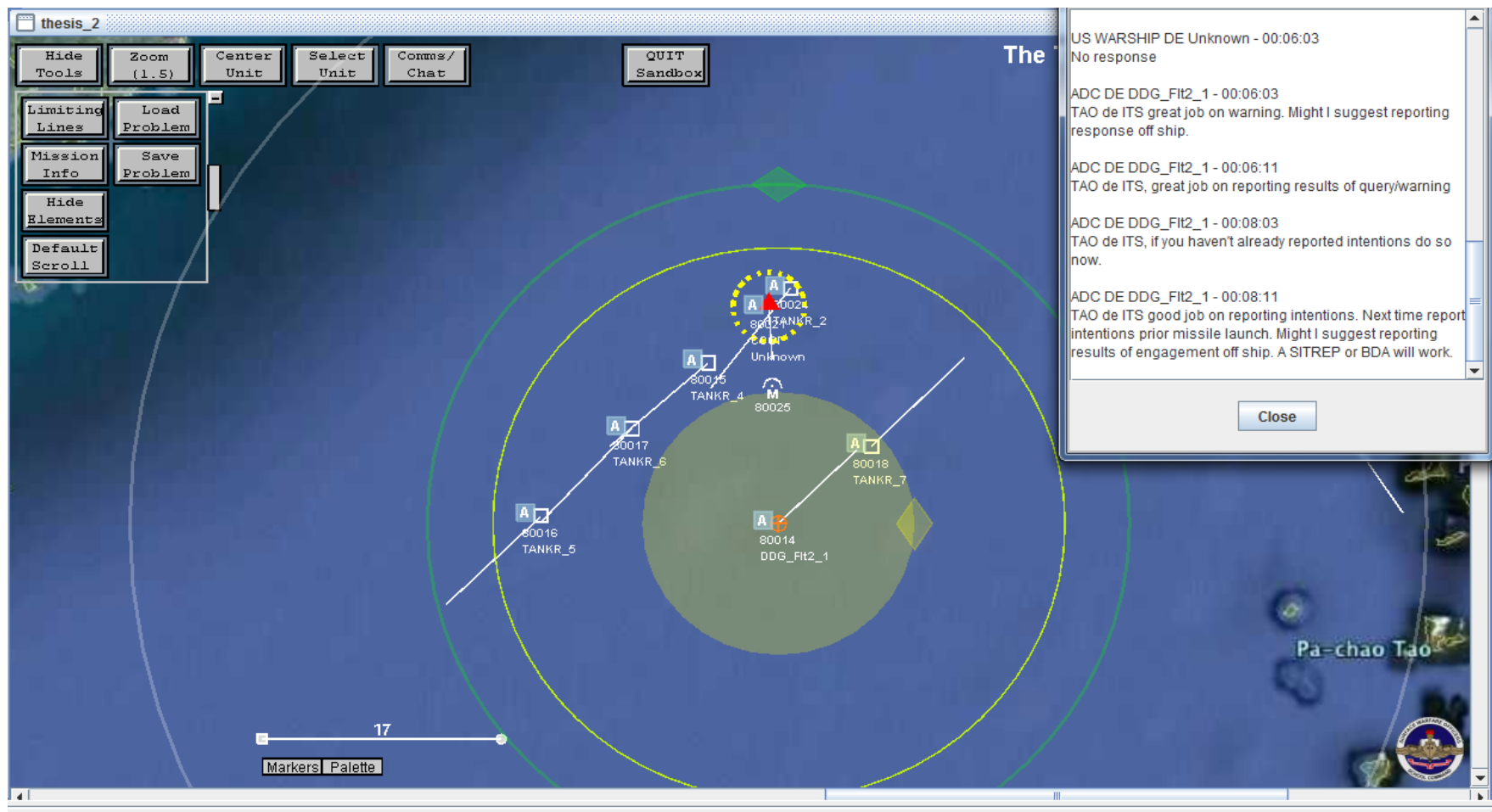


Figure 25. ITS suggesting user report intentions of engagement

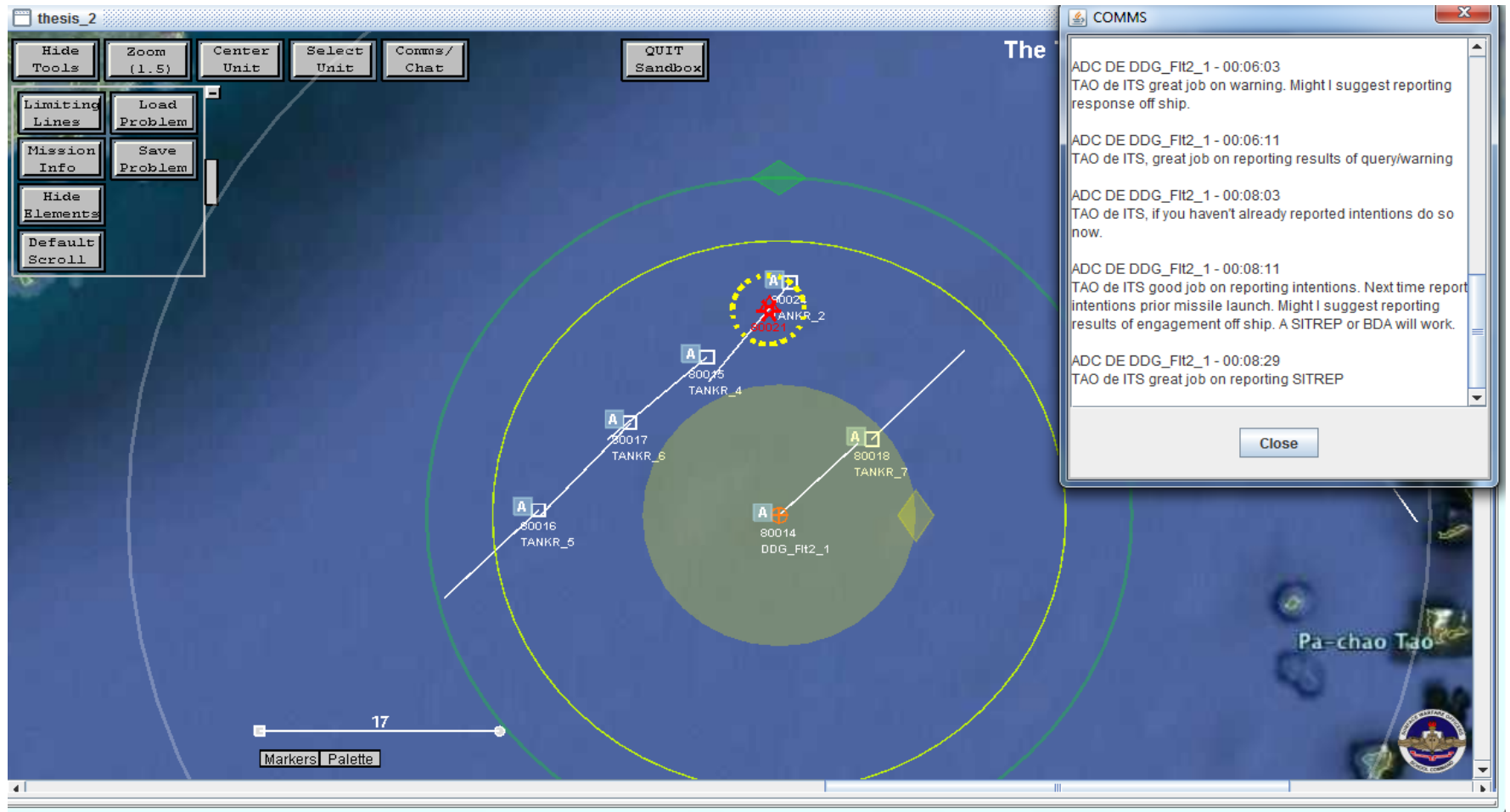
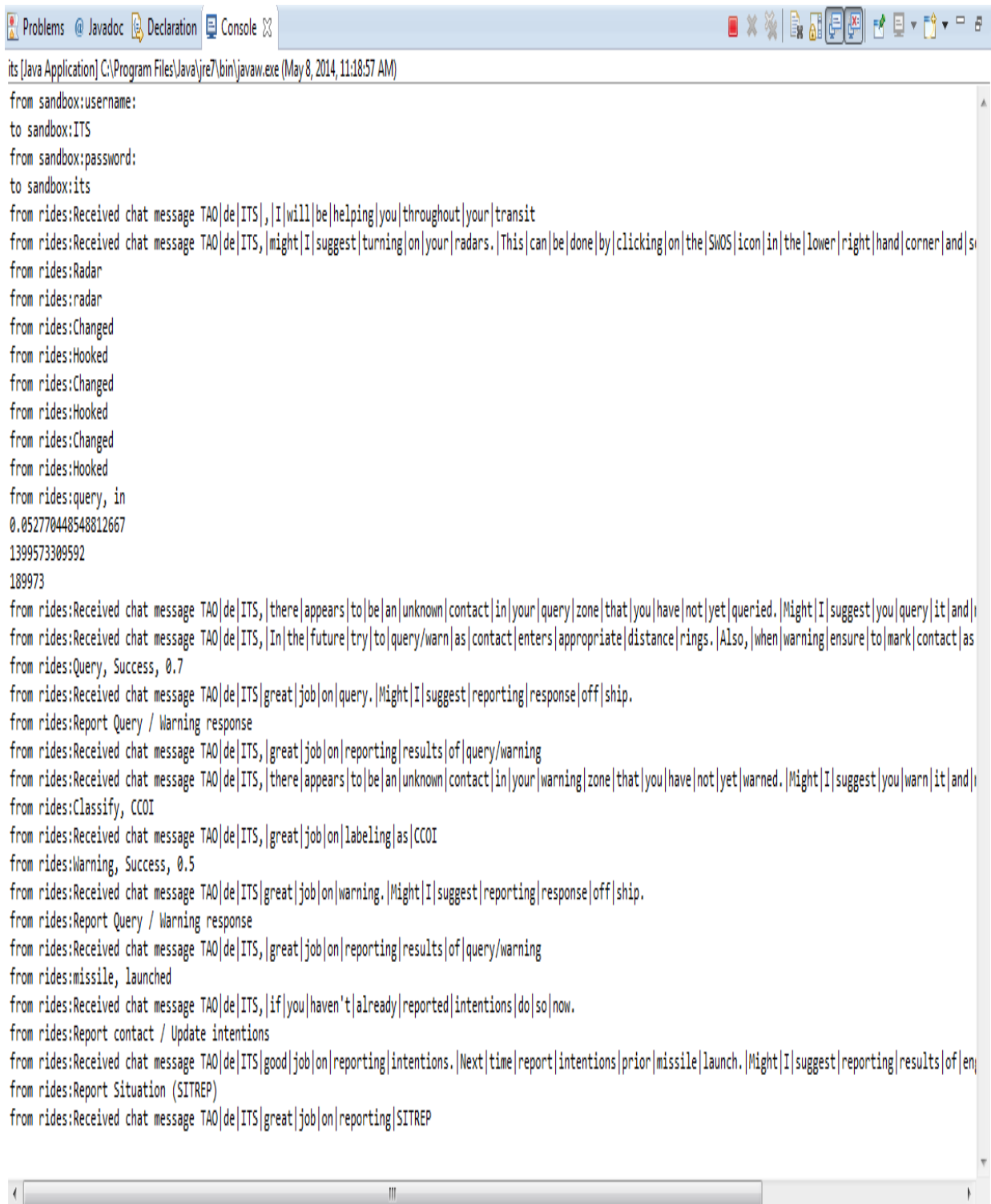


Figure 26. ITS providing feedback of successful engagement assessment



```
its [Java Application] C:\Program Files\Java\jre7\bin\javaw.exe (May 8, 2014, 11:18:57 AM)
from sandbox:username:
to sandbox:ITS
from sandbox:password:
to sandbox:its
from rides:Received chat message TAO/de|ITS|, I|will|be|helping|you|throughout|your|transit
from rides:Received chat message TAO/de|ITS|, might I|suggest|turning|on|your|radars. |This|can|be|done|by|clicking|on|the|SWOS|icon|in|the|lower|right|hand|corner|and|s
from rides:Radar
from rides:radar
from rides:Changed
from rides:Hooked
from rides:Changed
from rides:Hooked
from rides:Changed
from rides:Hooked
from rides:query, in
0.052770448548812667
1399573309592
189973
from rides:Received chat message TAO/de|ITS|, there|appears|to|be|an|unknown|contact|in|your|query|zone|that|you|have|not|yet|queried. |Might|I|suggest|you|query|it|and|
from rides:Received chat message TAO/de|ITS|, In|the|future|try|to|query|warn|as|contact|enters|appropriate|distance|rings. |Also, |when|warning|ensure|to|mark|contact|as
from rides:Query, Success, 0.7
from rides:Received chat message TAO/de|ITS|great|job|on|query. |Might|I|suggest|reporting|response|off|ship.
from rides:Report Query / Warning response
from rides:Received chat message TAO/de|ITS|, great|job|on|reporting|results|of|query|warning
from rides:Received chat message TAO/de|ITS|, there|appears|to|be|an|unknown|contact|in|your|warning|zone|that|you|have|not|yet|warned. |Might|I|suggest|you|warn|it|and|
from rides:Classify, CCOI
from rides:Received chat message TAO/de|ITS|, great|job|on|labeling|as|CCOI
from rides:Warning, Success, 0.5
from rides:Received chat message TAO/de|ITS|great|job|on|warning. |Might|I|suggest|reporting|response|off|ship.
from rides:Report Query / Warning response
from rides:Received chat message TAO/de|ITS|, great|job|on|reporting|results|of|query|warning
from rides:missile, launched
from rides:Received chat message TAO/de|ITS|, if|you|haven't|already|reported|intentions|do|so|now.
from rides:Report contact / Update intentions
from rides:Received chat message TAO/de|ITS|good|job|on|reporting|intentions. |Next|time|report|intentions|prior|missile|launch. |Might|I|suggest|reporting|results|of|en
from rides:Report Situation (SITREP)
from rides:Received chat message TAO/de|ITS|great|job|on|reporting|SITREP
```

Figure 27. Example of interaction between simulator and tutor

THIS PAGE INTENTIONALLY LEFT BLANK

IX. LESSONS LEARNED

In developing the ITS, I find it relevant to pass on some lessons learned that I struggled with early. These include hardware and software issues that may be relevant if a future student would like to continue with this work.

The computer that this was developed on was an early 2011 MacBook Pro. The single most limiting factor of this computer was that it had 4GB of Random Access Memory (RAM) stock. I would highly encourage anyone that intends to run a virtual machine, as is required to run TAO Sandbox on a Mac since it is Windows based, upgrade to at least 8GB. Using an application called Memory Clean I monitored the state of my RAM and often times found that with my 16GB upgrade that I would typically have 8GB free. Another upgrade that you might consider is a solid-state hard drive (SSD) or a hybrid hard drive disk (HHDD). Currently, the SSD is costly when going larger than 256GB but works very quickly. The HHDD is a good compromise between the speed of the SSD and the low cost for space of the conventional hard drive disk (HDD). I upgraded to a HHDD, which helps load the OS and virtual machine that runs windows quicker.

One major software issue that presented itself along the way was the choice of Integrated Development Environment (IDE). I was most familiar with Netbeans and chose to develop the ITS in this IDE. I found quickly that when running the ITS, via a TCP/IP connection, with the TAO Sandbox that it would often freeze, even after the memory upgrade, and would routinely go black. After many hours of troubleshooting, I was unable to remedy this and a MOVES instructor advised me to use Eclipse. Running Eclipse in the virtual machine worked without issue. Using another virtual machine, other than Parallels, might be an alternative solution as well but was not tested. This could present a future thesis as to what hardware/software configuration provides for optimum operating efficiency when running an external ITS to the TAO Sandbox. This may

be particularly necessary to do as the ITS becomes more robust and supports more warfare areas other than Air Defense.

X. CONCLUSIONS AND RECOMMENDATIONS

The experiment provided useful data that was able to demonstrate that statistically significant learning occurred when the experiment group was compared to the control group. However, this ITS is not ready for deployment as a training tool as there is much work to be done. I would like to note that many of the experimental users felt this could be a very useful tool for training combat watchstanders, or at the very least a tool that would allow for quality independent study to reinforce what is being taught at Department Head School. Both control and experimental users enjoyed the experience as it allowed them to refresh themselves in the air warfare domain and reminded them that skill degradation occurs quickly.

There are certain use cases that the expert model does not take into account. Several test subjects actions highlighted the need for these use cases in the experiment. First, there was no use case to correct the user if they queried or warned the aircraft prior to the aircraft entering into the respective ranges. While the advanced beginner or competent user groups and above might not need this case, the intention of an ITS is to be able to teach at all levels it is utilized.

A use case for utilizing a helicopter would be a welcome addition. Several participants attempted to utilize a helicopter to help visual identify the incoming contact as well as one case where the participant attempted to engage the contact with the helicopter. This takes the users actions a few steps beyond the intended scope of the thesis but would allow for continued improvement of this tutor.

Several users recommended an introduction video to TAO Sandbox that could demonstrate the actions available to them in the Sandbox or at the very least an interactive help function. This interactive help function could be

implemented into the tutor itself by allowing the users to ask how to complete certain tasks with the answers coded into the tutor being presented to the user.

Furthermore several students requested improvements to the simulator itself that would help improve graphical representation of the status of contacts such as marking the contacts as a COI or CCOI, as well as the capability to hook more than one target at a time. Others suggested voice communications being built into the system, as they did not feel the communications piece was realistic.

A. FUTURE WORK

As mentioned earlier in this chapter, this tutor is not ready for prime time. There is still much development and testing that must occur. Below are some suggestions on how to improve upon my work:

1. EEGs is a physiological measure that is a major target of researchers to help improve ITS assessment and remediation by evaluating what the subject was actually thinking when making decisions (Pokorny, 2012). Another physiological measure that has been incorporated in several studies, here at NPS, is eye tracking. Incorporate these two physiological measures into the analysis of this particular ITS and reevaluate the effectiveness of this tutor based on more than just metrics of performance.
2. There are several noted shortfalls of this ITS. Correct these and readdress the effectiveness of the tutor. I have shown an increase in scores, but it's possible with additional cases in the code to reach more level of students and increase scores even more. Personal observations showed that several of the subjects were doing some actions prior to that of when an expert would do so. Doing these actions caused a boolean to become true and would thus not provide assistance; when in fact assistance was warranted to notify the subject they were going a tad bit earlier.
3. One subject noticed that the query and warnings did not take into account law of the sea. This is largely due in part to the fact that the simulator, TAO Sandbox, does not take into account terrain. It would be beneficial to provide the ability to the simulator to recognize terrain as well as assisting the simulator in presenting a better understanding of territorial limitations presented by law of the sea.

4. This experiment was done with a limited number of participants consisting of a mixture of surface warfare officers and MOVES students. It would be a worthwhile venture to incorporate this experiment on a larger group to evaluate whether the gains were as significant. Furthermore, I would suggest taking this to SWOS and incorporating into ASAT or DHS. These groups would allow for a larger group that was versed in air warfare as well as participants that are currently studying this warfare area. It is truly amazing how quickly skills diminish in such a short amount of time.
5. I have mentioned that skills can diminish rapidly over a short amount of time and this experiment was conducted in a little less than an hour and a half. The results were significant even in a short-term memory recall situation. It would be interesting to study the results of a long-term memory recall. A study conducted over a period of months might demonstrate this. It might be useful to conduct the same experiment; however, placing a time of about a month in between subjects being exposed to the scenario to see if the tutor has any effect on long term memory recall.
6. This ITS was built for an air warfare domain in a limited scope. Broadening the scope to include multiple hostile contacts and/or situations where the user has to make a decision as to whether a target should be engaged would be beneficial to improving the ability of this tutor in training surface warfare officers in air warfare.
7. Further development of this tutor into other warfare areas would also be quite beneficial. Air warfare is one of many warfare areas that could benefit from standardized training and the ability for students to take an instructor anywhere with them in a PORTS type configuration. Further develop this ITS to handle other warfare areas such as ASW, MIW, and SW.
8. I cautioned several times throughout that this tutor has not been validated nor verified. Have the expert system validated by experts in the field of air warfare.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. ITS CODE

```
package its;

import java.io.*;
import java.net.*;
import java.util.Calendar;

/**
 *
 * @author Ryan McBride
 */
public class its {

    /**
     * @param args the command line arguments
     * @throws InterruptedException
     */
    public static void main(String[] args) throws IOException, InterruptedException {
        Socket socket = null;
        PrintWriter out_to_sandbox = null;
        BufferedReader in_from_sandbox = null;
        boolean halt = false;
        boolean radarOn = false;
        boolean radarNotification = false;
        boolean hooked = false;
        boolean hookNotification = false;
        boolean changed = false;
        boolean changeNotification = false;
        boolean vitalPrompt = false;
        boolean warningPrompt = false;
        boolean inWarningZone = false;
        boolean missileLaunched = false;
        boolean queryPresented = false;
        boolean warningPresented = false;
        boolean queryPrompt = false;
        Calendar cal = Calendar.getInstance();
        long time = cal.getTimeInMillis();
        boolean queryTimeSet = false;
        long inQuery = 0;
        long maxQueryTime = 30000;
        long maxWarningTime = 15000;
        long estInWarning = 0;
        double queryRange = 45;
        double warningRange = 25;
        double enemySpeed = 379;
        int count = 0;

        try {
```

```

//connecting to server on localhost, port 5000; These values are defined in
btl.prop

socket = new Socket("localhost", 5000);
out_to_sandbox = new PrintWriter(socket.getOutputStream(), true);
in_from_sandbox = new BufferedReader(new InputStreamReader(
    socket.getInputStream()));
} catch (UnknownHostException e) {
    System.err.println("Don't know about host: localhost.");
    System.exit(1);
} catch (IOException e) {
    System.err.println("Couldn't get I/O for "
        + "the connection to: localhost.");
    System.exit(1);
}
String sandboxInput;
//exchanging username
sandboxInput = in_from_sandbox.readLine();
System.out.println("from sandbox:"+sandboxInput);
if (sandboxInput.equals("username:")){
    System.out.println("to sandbox:"+ "ITS");
    out_to_sandbox.println("ITS");
}
//exchanging password
sandboxInput = in_from_sandbox.readLine();
System.out.println("from sandbox:"+sandboxInput);
if (sandboxInput.equals("password:")){
    System.out.println("to sandbox:"+ "its");
    out_to_sandbox.println("its");
}
out_to_sandbox.println("chat                                     message
TAO|de|ITS|,|I|will|be|helping|you|throughout"
    + "|your|transit");
BufferedReader stdIn = new BufferedReader(new InputStreamReader(System.in));
//String userInput;

//exchanging data with the Sandbox; whatever's typed into term window is sent to
the

//Sandbox and whatever is send from the Sandbox is written to the term window
while (stdIn.ready() || in_from_sandbox.ready() || !halt){
    //Gives student one minute to turn radars on before notifying them
    if(Calendar.getInstance().getTimeInMillis() > (time + 60000) && !radarOn
&&
        !radarNotification){
            out_to_sandbox.println("chat                                     message
TAO|de|ITS|,|might|I|suggest|turning"
                + "|on|your|radars."
                +
                "|This|can|be|done|by|clicking|on|the|SWOS|icon|in|the|lower"
                + "|right|hand|corner|and"
                + "|selecting|turn|on|all|radars.");
            radarNotification = true;

```

```

    }
    //Gives student 2 minutes to hook contact
    if(Calendar.getInstance().getTimeInMillis() > (time + 150000) && !hooked
&&
        !hookNotification){
            out_to_sandbox.println("chat
message
TAO|de|ITS,|I|noticed|you|havent|hooked"
            + "|any|contacts.|This|can"
            +
            "|be|done|by|clicking|on|the|A|next|to|the|target|you|would|like|to"
            + "|hook,|and|selecting|hook."
            +
            "|This|will|also|display|the|kinematics|of|the|target|to|include|range,"
            + "|bearing,|speed,|course,"
            + "|and|altitude.");
            hookNotification = true;
            //count++;
        }
        if(Calendar.getInstance().getTimeInMillis() > (time + 170000) && !changed
&& !changeNotification && count < 2){
            out_to_sandbox.println("chat
message
TAO|de|ITS,|I|noticed|you|aren't|changing|"
            + "contacts|of|interest."
            +
            "|Might|I|suggest|you|evaluate|all|contacts|on|radar|plot.");
        }
        //System.out.println(queryPresented+", "+queryTimeSet+", "+
(Calendar.getInstance().getTimeInMillis() > (inQuery + maxQueryTime)));
        if(!queryPresented && queryTimeSet && !queryPrompt &&
            (Calendar.getInstance().getTimeInMillis() > (inQuery +
maxQueryTime))){
            out_to_sandbox.println("chat
message
TAO|de|ITS,|there|appears|to|be|an"
            +
            "|unknown|contact|in|your|query|zone|that|you|have|not|yet"
            +
            "|queried.|Might|I|suggest|you|query|it|and|mark|it|as|a|COI.");
            Thread.sleep(1000);
            out_to_sandbox.println("chat
message
TAO|de|ITS,|In|the|future|try"
            +
            "|to|query/warn|as|contact|enters|appropriate|distance|rings."
            +
            "|Also,|when|warning|ensure|to|mark|contact|as|a|CCOI.");
            queryPrompt = true;
        }
        if(!warningPresented && queryTimeSet && !warningPrompt &&
            (Calendar.getInstance().getTimeInMillis() > (estInWarning
+ maxWarningTime))){

```

```

        out_to_sandbox.println("chat message
TAO|de|ITS,|there|appears|to|be|an"
+
"|unknown|contact|in|your|warning|zone|that|you|have|not|yet"
+
"|warned.|Might|I|suggest|you|warn|it|and|mark|it|as|a|CCOI.");
        warningPrompt = true;
    }
    //System.out.println("Looking for output from sandbox");
    if (in_from_sandbox.ready()){
        //reading data from the Sandbox
        sandboxInput = in_from_sandbox.readLine();
        if (sandboxInput != null && !sandboxInput.equals("null")){
            if (sandboxInput.length() > 0){
                //writting data from the Sandbox to term window
                System.out.println("from rides:"+sandboxInput);
                if (sandboxInput.contains(", ")){
                    String[] fromSandbox =
sandboxInput.split(", ");

                    if(fromSandbox[0].equalsIgnoreCase("classify")){

                        if(fromSandbox[1].equalsIgnoreCase("coi")){

                            out_to_sandbox.println("chat message TAO|de|ITS,|great|job|on|"
+
"labeling|as|COI");

                                }
                                else
                                if(fromSandbox[1].equalsIgnoreCase("ccoi")){

                                    out_to_sandbox.println("chat message TAO|de|ITS,|great|job|on|"
+
"labeling|as|CCOI");

                                        }
                                        }
                                        /*if(!labelCOI && inQueryZone &&
!COIwarning){
                                            out_to_sandbox.println("chat
message TAO|de|ITS,|there|is|an|unknown
|contact|in|query|zone.|Might|I|suggest|to|label|as|COI.");
                                            COIwarning = true;
                                        }
                                        if(!labelCCOI && inWarningZone &&
!CCOIwarning){
                                            out_to_sandbox.println("chat
message TAO|de|ITS,|there|is|an|unknown
|contact|in|warning|zone.|Might|I|suggest|to|label|as|CCOI.");
                                            CCOIwarning = true;

```

```

        }*/
        if
        (fromSandbox[0].equalsIgnoreCase("query")){

            if(fromSandbox[1].equalsIgnoreCase("success") && !inWarningZone){

                out_to_sandbox.println("chat message TAO|de|ITS|great|job|on|query."
                                     +
                                     "|Might|I|suggest|reporting|response|off|ship.");

                queryPresented = true;
            }
            else
            if(fromSandbox[1].equalsIgnoreCase("in")){

                inQuery =
                Calendar.getInstance().getTimeInMillis();

                queryTimeSet = true;
                //dividing by 3600000
                //queryRange -
                converts from hours to milliseconds
                warningRange/enemySpeed is applying the distance formula D=ST
                //distance = speed*time
                or rearranged time = distance/speed
                long DST = (long)
                (((queryRange-warningRange)/enemySpeed)*3600000);

                System.out.println((queryRange-warningRange)/enemySpeed);

                System.out.println(inQuery);

                System.out.println(DST);

                //adding 15 seconds for
                user delay and conversion error
                estInWarning = inQuery
                + DST + 15000;

                //inQueryZone = true;
            }*/
            //this code was intended for
            //only one watcher is
            implemented
            else
            if(fromSandbox[1].equalsIgnoreCase("warning")){

                out_to_sandbox.println("chat message TAO|de|ITS|you|have|not|yet|
                queried|the|aircraft.|Recommend|doing|so|now.");
            }
            else
            if
            (fromSandbox[1].equalsIgnoreCase("failure")){

                out_to_sandbox.println("chat message TAO|de|ITS|you|failed|to|query,

```

```

        |make|sure|that|you|warn|incoming|contact|,|if|necessary");
                                                    //warningPrompt = true;
                                                    inWarningZone = true;
                                                    }
                                                    */
        }
        else
if(fromSandbox[o].equalsIgnoreCase("missile")){
                                                    out_to_sandbox.println("chat
message TAO|de|ITS,|if|you|haven't|already|"
                                                    +
"reported|intentions|do|so|now.");
                                                    missileLaunched = true;
        }
        else
        if
        (fromSandbox[o].equalsIgnoreCase("warning")){
            if(fromSandbox[1].equalsIgnoreCase("success") && !vitalPrompt){
                out_to_sandbox.println("chat message TAO|de|ITS|great|job|on|warning."
                +
                "|Might|I|suggest|reporting|response|off|ship.");
                warningPresented =
true;
            }
            else
            if
            (fromSandbox[1].equalsIgnoreCase("failure")){
                out_to_sandbox.println("chat message TAO|de|ITS|you|failed|to|warn|the|"
                +
                "incoming|aircraft|in|time.|The|aircraft|is|now|in|your|vital"
                +
                "|area.|Looks|to|be|hostile.|Might|I|suggest|engaging|the|aircraft.");
                vitalPrompt = true;
            }
        }
        }
        else if (!sandboxInput.contains(",")){
            String[] fromSandbox =
sandboxInput.split(" ");
            if
            (fromSandbox[o].equalsIgnoreCase("destroyed")){
                halt = true;
            }
            if(fromSandbox[o].equalsIgnoreCase("radar")){
                radarOn = true;
            }
        }
    }
}

```

```

        if(fromSandbox[o].equalsIgnoreCase("hooked")){
                                                    hookNotification = true;
        }

        if(fromSandbox[o].equalsIgnoreCase("changed")){
                                                    changeNotification = true;
                                                    count++;
        }
        if
(fromSandbox[o].equalsIgnoreCase("report")){

        if(fromSandbox[1].equalsIgnoreCase("query")){

        //out_to_sandbox.println("chat message TAO|de|ITS,|I|will|
        //be|helping|you|throughout|your|transit");

                                                    Thread.sleep(100);

        out_to_sandbox.println("chat message TAO|de|ITS,|great|job"
+
        "|on|reporting|results|of|query/warning");
                                                    }
                                                    else
        if(fromSandbox[1].equalsIgnoreCase("BDA")){

                                                    Thread.sleep(100);

        out_to_sandbox.println("chat message TAO|de|ITS|great|job"
+
        "|on|reporting|BDA");
                                                    }
                                                    else
        if(fromSandbox[1].equalsIgnoreCase("situation")){

                                                    Thread.sleep(100);

        out_to_sandbox.println("chat message TAO|de|ITS|great|job|on"
+
        "|reporting|SITREP");
                                                    }
                                                    else
        if(fromSandbox[1].equalsIgnoreCase("failure")){

                                                    Thread.sleep(100);

        out_to_sandbox.println("chat message TAO|de|ITS,|Don't|"
+
        "forget|to|report|query/warnings|results|and|intentions");
                                                    }
                                                    else if(fromSandbox.length >4
&& fromSandbox[4].equalsIgnoreCase("intentions") &&

```

```

        !missileLaunched){
                                                    Thread.sleep(100);

        out_to_sandbox.println("chat message TAO|de|ITS|great|job|on"
                                                    +
        "|reporting|intentions.|Don't|forget|to|report|BDA|"
                                                    +
        "or|SITREP|following|engagement.");
                                                    }
        && fromSandbox[4].equalsIgnoreCase("intentions") &&
                                                    else if( fromSandbox.length >4

        missileLaunched){
                                                    Thread.sleep(100);

        out_to_sandbox.println("chat message TAO|de|ITS|good|job|on"
                                                    +
        "|reporting|intentions.|Next|time|report|intentions"
                                                    +
        "|prior|missile|launch.|Might|I|suggest|reporting|"
                                                    +
        "results|of|engagement|off|ship.|A|SITREP|or|BDA|will|work.");
                                                    }
                                                    else{
                                                    }
                                                    }
        }else{
            halt = true;
        }
    }
    //this is in case the user wants to write commands into the terminal
window
    /*else if (stdIn.ready()){
    //reading data from term window
    userInput = stdIn.readLine();
    if (userInput != null){
        //writing data to the Sandbox
        out_to_sandbox.println(userInput);
    }else {
        halt = true;
    }
    }*/
}

//closing streams and socket
out_to_sandbox.close();
in_from_sandbox.close();

```

```
        socket.close();  
        System.out.println("connection closed");  
    }  
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. RECORD OF CHANGES

The ITS could not be utilized with TAO Sandbox without a few changes to the code of the TAO Sandbox. Below is a list of changes that were made:

Added

remote.socket.login.usernames.passwords.atts=ITS:its:.k.MsgTo_ITS:.k.MsgFrom_ITS/ to battle.prp

Added to .k.commMgr.init

```
.k.MsgFrom_ITS:=""
```

```
.k.MsgTo_ITS:=""
```

```
ITSactive:=true;
```

Added to .k.commMgr. an event called MessageFromITS with trigger

```
.k.MsgFrom_ITS != ""
```

Body being:

```
{  
  if LCase (GetField(.k.MsgFrom_ITS,1)) != "signoff"  
    ITSactive := true;  
  problemdir := concat (.sys.problemDir,"ITS/");  
  recorddir := concat (.sys.recordDir,"ITS/");  
  $str := .k.MsgFrom_ITS;  
  .k.MsgFrom_ITS:= "";  
  DoEvent (MessageReceived.,$str);  
}
```

Added attribute to k.commMgr. ITSactive with value of false

Added to .k attributes MsgTo_ITS and MsgFrom_ITS

//Providing comments to ITS

Added

Print ("Great job on query");

.k.MsgTo_ITS := "Query, Succes, 0.9";

To .k.theLine.Query_Finish

.k.watchersBtn inUse attribute set to true. Must set back to false when done.

For Revisiting Tracks

Modified to print execution line

.k.kinematics. ChangeUnit

.k.Watchr_copy.MakeLimitsReport

For BDA

Added .k.MsgTo_ITS := guival2; to .k.commsBtn.saveMessage

Added .k.MsgTo_ITS := "Radar"; to .k.SWOS.TurnOn AllRadar

Added .k.MsgTo_ITS:= "radar"; to .k.unit_master.RadarChange_air

Added .k. MsgTo_ITS := "Hooked"; to .k.hookInd.NewHook

Added.k.MsgTo_ITS := "Changed"; to .k.kinematics.ChangeUnit

Added .k.MsgTo_ITS:= "Classify, COI"; to .k.unit_master.Exec_A_D

Added .k.MsgTo.ITS:"Classify, CCOI"; to .k.unit.master.Exec_A_D

Added .k.MsgTo_ITS := "query, in"; to .k.Watchr5.MakeLimits Report

Added .k.MsgTo_ITS := "warning, failure"; to .k.Watchr.MakeLimitsReport

Added .k.MsgTo_ITS := "Destroyed"; to .k.ship_master.Destroy

Added .k.MsgTo_ITS := "report failure"; to
.k.commsBtn.QueryWarningReportWatcher

Added a conditional (or .k.commMgr.ITSactive) to .k.commMgr.UserLoggedIn

Added a conditional (or .^commMgr.ITSactive) to chatWindow.record

Added ITS variable to sign on an off of .k.commMgr

Added if statement to .k.playbackStartRecording

```
if (.k.commMgr.ITSactive){  
    .k.msg_to_ITS = $newString;  
}
```

Added message lines to .k.theLine.QueryFinish

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. PRE-SCENARIO SURVEY

Before Scenario Survey

Are you a qualified Tactical Action Officer (TAO)?	YES	NO
How long have you stood watch as a TAO? (Estimate in hours)		
Are you familiar with the Detect to Engage sequence?	YES	NO
Have you used the TAO Sandbox before?	YES	NO

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. POST-SCENARIO SURVEY

After Scenario Survey

Was the scenario realistic? (If no, please provide comments below to improve scenario)	YES	NO
Do you feel the tutor gave appropriate feedback to your actions? (If no, please provide comments below to help improve the tutor)	YES	NO
How would you improve the tutor? (Aside from comments about feedback i.e. voice response, more or less feedback)		

General

Comments:

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins or other markings on the paper.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E. NAVY SIMULATORS

SOPHIE was developed to teach problem-solving skills by having the student take measurements on a simulated electronic piece of equipment, which has a malfunction. The student's goal was to find the fault by troubleshooting the simulated circuit. SOPHIE was one of the earliest version of an ICAI and provided much insight to the field as well as a new approach to training Sailors. More specifically, "SOPHIE'S contribution to the ICAI field is the introduction of device-based simulation to support checking of student inferences, as well as heuristic strategies to allow question generation and answering mechanisms" (Salgado-Zapata, 1989).

STEAMERs main goal was to replicate a ship's steam propulsion plant and to train operators by helping them understand the complicated plant through the use of graphical user interfaces (GUI).

QUEST, is very similar to SOPHIE, in that it was developed to help technicians troubleshoot circuit problems. The difference is that it used graphic simulation in its approach to train Sailors.

IMTS also was used to train Sailors in their systems. IMTS was used in the simulation of the SH-3 helicopter's blade folding mechanism. IMTS may have been one of the earliest systems that attempted to measure the student's knowledge in order to select its tutoring strategy.

MACH-III provided training in the maintenance of electronic and electromechanical systems particularly the maintenance of the illumination radar of the HAWK air defense system. MACH-III supported three modes of instruction to include demonstration, step-by-step guided practice, and free-form monitored practice.

TRIO trained F-14 pilots and radar officers in air-intercept problems. TRIO is a more advanced ICAI, similar in idea to the ITS I developed for the TAO Sandbox. The advancement of TRIO incorporated real-time simulation, speech synthesis, and speech recognition capabilities. With TRIO, the student's solution is compared to the solution developed by an expert knowledge base system. The expert knowledge system is the optimal solution given the parameters of the situation. The advancement of TRIO centered ICAI on not only teaching students how to correctly solve a problem, but to do so in a manner fast enough to react to radar warnings it presented the student. This added an element of stress, which is often times, experienced in a real world-operating environment.

INCOFT, also a time-constrained application, was intended to train students in the operation of the engagement control station of a PATRIOT air defense missile system. INCOFT used the same approach as TRIO, with exception of the speech capabilities (Salgado-Zapata, 1989).

More recently there has been an influx of simulators, some of which have had ITSs developed for them. These include Navigation Seamanship Shiphandler Trainer (NSST), Conning Officer Virtual Environment (COVE), Battle Force Tactical Trainer (BFTT), Multi-Mission Tactical Trainer (MMTT), Tactical Action Officer (TAO) Sandbox, Tactical Action Officer/Intelligent Tutoring System (TAO/ITS), Full Mission Bridge (FMB), Littoral Combat Ship (LCS), Firearms Training Simulator Systems (FATS), and Engineering Trainers. All of these have proved invaluable in the training of officers, particularly as budget constraints are limiting real-world training opportunities.

The NSST program brings cutting edge technology to the fleet and delivers a valuable, high fidelity ship handling trainer, surpassing that of past simulation resources. NSST also provides increased simulator availability for ships' bridge teams by offering facilities at all Fleet Concentration Areas (FCAs). The purpose of NSST is to provide a synthetic environment where the ships' officers and enlisted personnel ship handling proficiency can be maintained while leveraging the ship handling talents of former commanders at sea and to provide

robust training that meets Standards of Training, Certification and Watchkeeping for Seafarers (STCW) equivalency requirements and creates more well-rounded officers and enlisted personnel (Gelinne, 2010).

Conning Officer Virtual Environment (COVE) stations provide state of the art navigation and shiphandling training for all of our Surface Officer's. These trainers can emulate every one of the U.S. Navy's homeports in addition to almost every routine port of call around the world. There are two types of COVE stations. Our COVE I stations are used primarily for our ASAT students. These stations consist of a virtual reality (VR) helmet that gives them a 360-degree view of their surroundings. Using a state of the art voice recognition system, students can give commands to the virtual helmsman, which are repeated back by the computer. Cove 3 stations have the same functionality as COVE I stations except they are viewed on three 50 inch displays to allow the student a wider field of view (FOV). Coupled with the VR helmet, the student has a 360 degree view. Department Head, Prospective Commanding Officers/ Prospective Executive Officer (PCO/PXO) and Major Command students primarily use the COVE 3 stations (Surface Warfare Officers School Command, 2013).

The BFTT family of systems (BFTT, BFTT Electronic Warfare Trainer (BEWT) and the Trainer Simulator Stimulator System (TSSS)) provides coordinated stimulation/simulation of shipboard combat systems to facilitate Combat Systems Team training providing the capability to conduct realistic joint warfare training across the spectrum of armed conflict and conduct realistic unit level team training in all primary warfare areas. BFTT accomplishes this by establishing a synthetic environment in which a tactical scenario is run. Combat System elements receive information from BFTT to stimulate shipboard tactical equipment, resulting in effective coordinated team training even while in-port. BFTT's ability to simulate friendly, neutral, or enemy forces such as aircraft/ships/submarines/ weapons allows for robust scenario development accurately reflecting the wide range of tactical proficiency levels as a ship and battle group traverse the deployment training cycle. In addition to this capability

to conduct single ship organic training, BFTT, when networked with other units through the Navy Continuous Training Environment (NCTE) allows for multiple units to participate in coordinated distributed training events. More than 100 U.S. Navy warships, including CVN, CG, DDG, LHA, LHD, LPD and LSD class ships have the BFTT system (United States Navy, 2013).

The Multi-Mission Tactical Trainer is a PC (Personal Computer) based Open Architecture Training System (PORTS), which means we can configure the system in multiple ways to simulate multiple platforms, so that students can optimize their training. The MMTT is mainly utilized by Department Head students, but can be utilized by all students to practice their tactical proficiency in multiple warfare areas, through the use of computer simulation. This allows students to practice what they have learned through classroom lectures and study. In Department Head School, the students are mentored as they progress from a basic scenario up to a highly complex multi-warfare scenario to prepare them as Tactical Action Officer's. The Multi-Mission Tactical Trainer is one of SWOS' highly utilized trainers and its value is immeasurable (Surface Warfare Officers School Command, 2013).

TAO Sandbox, which happens to be the focal simulator for this thesis, is a PC based decision-making tool for tactical planning. TAO Sandbox was developed through a joint venture between Surface Warfare Officer School (SWOS), Office of Naval Research (ONR) and University of Southern California (USC), in order to provide a visual perspective to help students in planning tactical missions. By using the tools provided in Sandbox, the student can plan force movements, anticipate enemy movements and evaluate responses through a visual medium without putting sailors or ships in any danger. Sandbox is a program that is rapidly developing and has developed from a basic visual aid to an invaluable support tool within the classroom environment (Surface Warfare Officers School Command, 2013).

The Tactical Action Officer/Intelligent Tutoring System is one of SWOS' newest tools in training and developing the Navy's tactical action officer (TAO) students. Using voice recognition software, the student plays the role of a TAO onboard one of the Navy's newest ships. Previously, it would take upwards of eight instructors simulating various watchstander roles to simulate the realism of leading a combat team in a ship's CIC. TAO/ITS eliminates that workload and the computer acts as all the applicable watchstanders. This allows one instructor to mentor 30 students at a time. The ITS is designed to tutor students on warfare areas where they may be weak. For example, if a student does not have a lot of experience in air defense, the ITS can be used to help that student gain experience in this warfare area. The ITS allows the students to select the warfare area they would like to practice as well as levels from novice to advanced. Furthermore, the SWOS staff has the ability to script each scenario, making them basic for beginner students to highly complex for the more advanced student. Because the capability to author scenarios resides "in-house" SWOS can keep up to date on worldwide events, allowing the student to train anywhere in the world, virtually (Surface Warfare Officers School Command, 2013). This fidelity adds a sense of realism to the user not previously attainable.

The Littoral Combat Ship Full Mission Bridge is based on one of the Navy's newest classes of ships. It is a full sized trainer that students can train on in preparation of reporting to an LCS 1 class ship. Using the same software as FMB and COVE, the LCS trainer has every Navy homeport modeled and allows the student to navigate in and out of designated ports using the highly sophisticated controls of a real LCS. LCS has the capability to integrate with FMB and COVE 3 stations to allow for multi-platform scenarios (Surface Warfare Officers School Command, 2013).

The F.A.T.S. is part of the Level III pre-command AT Training requirements specified in DoD INST 2000.16 and implemented by CENSECFOR in the Commanding Officers Anti-Terrorism (COAT) course. This simulator uses

wireless technology and rechargeable magazines to allow students, armed with a 9mm and M-4, to face possible lethal threat situations and not only have to make instant decisions, but know if it was the right decision and who survived the encounter. The goal of this training is to give students judgmental training to include target discrimination, force escalation/de-escalation, and individual leadership imperatives (Surface Warfare Officers School Command, 2013).

Engineering students at SWOS are also using technology to their advantage. CG47 Smart Ship Controls, FFG 7, DDG 51, MCM 1, LSD 41 and LCS ship class students start off using PC based trainers for their applicable class. These trainers simulate steam plants and diesel gas turbine plants. This PC based training prepares the students to make the transition to full size trainers. It is in these full size trainers that students get valuable hands-on training. By the end of the course, assessors from the Navy's Afloat Training Group Engineering Assessments commands will assess students upon successful completion of the course. This streamlines the process of the student qualifying as an Engineer Officer of the Watch (EOOW) when they reach their ship. The LCS RCO console is one of N74 Directorate's newer trainers. It is designed to train Engineering Department Heads, Main Propulsion Assistants (MPA), and Senior Enlisted RCOs assigned to LCS 1 class ship, which employs Smart Ship Technology. This trainer simulates the bridge watch station a RCO would operate on a LCS 1 class ship. Using touch screen technology, the student can manipulate the plant as they would on a LCS 1 class ship (Surface Warfare Officers School Command, 2013).

APPENDIX F. DATA EXTRACTED FROM TAO SANDBOX

Control		Run 1	Run 3	Delta	Avg Score -All Runs
Subject 1	NTA1	0.95	0.9	-0.05	0.91
	NTA2	0	0.88	0.88	0.88
	NTA4	0.45	0.2	-0.25	0.4
	NTA5	0.1	0.1	0	0.1
	NTA6	0.24	0.24	0	0.24
	NTA7	0.05	0.05	0	0.05
Subject 2	NTA1	0.7	0.95	0.25	0.91
	NTA2	0	0	0	0
	NTA4	0.23	0	-0.23	0.35
	NTA5	0.23	0.1	-0.13	0.18
	NTA6	0.9	0.9	0	0.86
	NTA7	0	0.325	0.325	0.42
Subject 3	NTA1	0.95	0.95	0	0.95
	NTA2	0	0	0	0
	NTA4	0.53	0	-0.53	0.53
	NTA5	0.425	0.1	-0.33	0.26
	NTA6	0.68	0.74	0.06	0.77
	NTA7	0.05	0.9	0.85	0.33
Subject 4	NTA1	0.95	0.95	0	0.95
	NTA2	0.88	0.88	0	0.88
	NTA4	0.45	0.2	-0.25	0.32
	NTA5	0.25	0.1	-0.15	0.17
	NTA6	0.9	0.57	-0.33	0.73
	NTA7	0	0.33	0.33	0.19
Subject 5	NTA1	0.83	0.95	0.12	0.91
	NTA2	0.88	0	-0.88	0.88
	NTA4	0.2	0	-0.2	0.3
	NTA5	0.17	0.1	-0.07	0.14
	NTA6	0	0.24	0.24	0.24
	NTA7	0	0	0	0

Table 2. Control Group Subjects 1-5

Control		Run 1	Run 3	Delta	Avg Score -All Runs
Subject 6	NTA1	0.95	0	-0.95	0.48
	NTA2	0.88	0.88	0	0.88
	NTA4	0.53	0	-0.53	0.27
	NTA5	0.43	0.1	-0.33	0.27
	NTA6	0.24	0	-0.24	0.12
	NTA7	0.05	0.05	0	0.05
Subject 7	NTA1	0.25	0.7	0.45	0.55
	NTA2	0.88	0.88	0	0.88
	NTA4	0.2	0.3	0.1	0.23
	NTA5	0.1	0.1	0	0.15
	NTA6	0	0	0	0.5
	NTA7	0.05	0.05	0	0.05
Subject 8	NTA1	0.95	0.95	0	0.92
	NTA2	0.88	0	-0.88	0.88
	NTA4	0.2	0.2	0	0.2
	NTA5	0.1	0.1	0	0.1
	NTA6	0	0	0	0
	NTA7	0.05	0.05	0	0.05
Subject 9	NTA1	0.95	0.95	0	0.95
	NTA2	0.88	0.88	0	0.88
	NTA4	0.2	0.2	0	0.2
	NTA5	0.1	0.1	0	0.1
	NTA6	0	0	0	0
	NTA7	0.05	0.05	0	0.05
Subject 10	NTA1	0.83	0.95	0.12	0.9
	NTA2	0.88	0.88	0	0.88
	NTA4	0.6	0	-0.6	0.6
	NTA5	0.48	0.1	-0.38	0.26
	NTA6	0.9	0.9	0	0.83
	NTA7	0.9	0.9	0	0.9
SUM		25.505	21.925	-3.58	
AVG NTA1		0.831	0.825	-0.01	0.843
AVG NTA2		0.616	0.528	-0.09	0.704
AVG NTA4		0.359	0.11	-0.25	0.34
AVG NTA5		0.2385	0.1	-0.14	0.173
AVG NTA6		0.386	0.359	-0.03	0.429
AVG NTA7		0.12	0.2705	0.151	0.209

Table 3. Control Group Subjects 6-10 and Averages

Experiment		Run 1	Run 3	Delta	Avg Score-All Runs
Subject 1	NTA1	0.83	0.95	0.12	0.93
	NTA2	0.88	0.88	0	0.88
	NTA4	0.2	0.7	0.5	0.4
	NTA5	0.1	0.1	0	0.2
	NTA6	0	0.9	0.9	0.64
	NTA7	0.05	0.9	0.85	0.33
Subject 2	NTA1	0.95	0.95	0	0.95
	NTA2	0	0	0	0
	NTA4	0.27	0.7	0.43	0.38
	NTA5	0.2	0.1	-0.1	0.15
	NTA6	0.9	0.9	0	0.77
	NTA7	0.9	0.9	0	0.62
Subject 3	NTA1	0.91	0.95	0.04	0.93
	NTA2	0	0.88	0.88	0.88
	NTA4	0.63	0.7	0.07	0.64
	NTA5	0.53	0.1	-0.43	0.31
	NTA6	0.57	0.9	0.33	0.68
	NTA7	0.6	0.6	0	0.42
Subject 4	NTA1	0.95	0.95	0	0.95
	NTA2	0.88	0.88	0	0.88
	NTA4	0.3	0.7	0.4	0.62
	NTA5	0.3	0.1	-0.2	0.2
	NTA6	0.9	0.9	0	0.81
	NTA7	0.6	0.9	0.3	0.4
Subject 5	NTA1	0	0.95	0.95	0.95
	NTA2	0	0	0	0
	NTA4	0.23	0.7	0.47	0.35
	NTA5	0.2	0.1	-0.1	0.15
	NTA6	0	0.9	0.9	0.9
	NTA7	0.9	0.9	0	0.39

Table 4. Experiment Group Subjects 1-5

Experiment		Run 1	Run 3	Delta	Avg Score-All Runs
Subject 6	NTA1	0.95	0.95	0	0.92
	NTA2	0.88	0.88	0	0.88
	NTA4	0.7	0.7	0	0.6
	NTA5	0.1	0.1	0	0.31
	NTA6	0	0.9	0.9	0.9
	NTA7	0.05	0.48	0.43	0.19
Subject 7	NTA1	0.95	0.95	0	0.95
	NTA2	0.88	0.88	0	0.88
	NTA4	0.2	0.9	0.7	0.4
	NTA5	0.2	0.1	-0.1	0.15
	NTA6	0.9	0.9	0	0.77
	NTA7	0	0.9	0.9	0.48
Subject 8	NTA1	0.9	0.95	0.05	0.93
	NTA2	0.88	0.88	0	0.88
	NTA4	0.6	0.9	0.3	0.6
	NTA5	0.43	0.1	-0.33	0.26
	NTA6	0.24	0.24	0	0.24
	NTA7	0.05	0.9	0.85	0.33
Subject 9	NTA1	0.83	0.95	0.12	0.93
	NTA2	0.88	0.88	0	0.88
	NTA4	0.23	0.23	0	0.23
	NTA5	0.17	0.1	-0.07	0.13
	NTA6	0	0.24	0.24	0.24
	NTA7	0.05	0.05	0	0.05
Subject 10	NTA1	0.83	0.95	0.12	0.93
	NTA2	0.88	0.88	0	0.88
	NTA4	0.2	0.7	0.5	0.37
	NTA5	0.1	0.1	0	0.2
	NTA6	0	0.24	0.24	0.24
	NTA7	0.05	0.05	0	0.05
SUM		26.91	38.07	11.16	
AVG NTA1		0.81	0.95	0.14	0.937
AVG NTA2		0.616	0.704	0.088	0.704
AVG NTA4		0.356	0.693	0.337	0.459
AVG NTA5		0.233	0.1	-0.133	0.206
AVG NTA6		0.351	0.702	0.351	0.619
AVG NTA7		0.325	0.658	0.333	0.326

Table 5. Experiment Group Subjects 6-10 and Averages

LIST OF REFERENCES

- Anderson, J. R., Boyle, C., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. In *Artificial Intelligence* (pp. 7–49). Elsevier Science Publishers B.V.
- Anderson, V. L., & McLean, R. A. (1974). *Design of Experiments: A Realistic Approach*. New York: Marcel Dekker.
- Armstrong, P. (2014). *Bloom's Taxonomy*. Retrieved May 1, 2014, from Vanderbilt University Center for Teaching: <http://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/>
- Baddeley, A., & Hitch, G. J. (2011, October 21). *Working memory*. Retrieved from Scholarpedia: http://www.scholarpedia.org/article/Working_memory
- Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.
- Columbia Center for New Media Teaching and Learning (CCNMTL). (2003). *Two-Sample T-Test*. Retrieved from QMSS e-Lessons: http://ccnmtl.columbia.edu/projects/qmss/the_ttest/twosample_ttest.html
- De Veaux, R. D., Velleman, P. F., & Bock, D. E. (2008). *Stats: data and models - 2nd ed*. Boston, MA, USA: Pearson Education.
- Defense Acquisition University. (2010, February 12). *Measure of performance*. Retrieved from Acquisition Community Connection: <https://acc.dau.mil/CommunityBrowser.aspx?id=204186>
- EduTech Wiki. (2007, April 10). *Nine events of instruction*. Retrieved from EduTech Wiki: http://edutechwiki.unige.ch/en/Nine_events_of_instruction
- Federation of American Scientists. (n.d.). *Weapon posture 1*. Retrieved May 6, 2014, from Battle Orders: <http://www.fas.org/man/dod-101/navy/docs/swos/cmd/fts/Fts3/tsld021.htm>
- Freedman, R. (2000, Fall). What is an intelligent tutoring system? *Intelligence*, 11 (3), pp. 15–16.
- Gagne, R. M., & Briggs, L. J. (1979). *Principles of instructional design* (2nd Edition ed.). New York: Holt, Rinehart, and Winston.

- Gates, R. (2011, April 19). Science and technology (S&T) priorities for FY13-17 planning. Retrieved from ACQWeb:
<http://www.acq.osd.mil/chieftechnologist/publications/docs/OSD%2002073-11.pdf>
- Gelinne, J. (2010, February 19). *COMNAVSURFOR INSTRUCTION 3505.1A*. Retrieved February 24, 2014, from NSSTraining.net:
<http://www.nsstraining.net/3505.1a.pdf>
- Greenert, J. W. (2012). *U.S. Navy program guide 2012*. Retrieved December 10, 2013, from Department of the Navy:
<http://www.navy.mil/navydata/policy/seapower/npg12/top-npg12.pdf>
- Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 30-43.
- Munro, Allen. (2013, December 20). TAO Sandbox v. 3.89 release notes.
- Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education* (10), 98–129.
- Nkambou, R., Bourdeau, J., & Psyche, V. (2010). Building intelligent tutoring systems: An overview. In *Advances in Intelligent Tutoring Systems* (pp. 361–375). Berlin, Germany: Springer-Verlag.
- Nwana, H. S. (1990). Intelligent tutoring systems: An overview. *Artificial Intelligence Review*, 251–277.
- Office of Naval Research. (2011, September 1). *Naval S&T strategic plan*. Retrieved from Office of Naval Research: <http://www.onr.navy.mil/About-ONR/~media/Files/About-ONR/Naval-Strategic-Plan.ashx>
- Pearson Education, Inc. (2014). *About working memory*. Retrieved June 2, 2014, from COGMED: <http://www.cogmed.com/about-working-memory>
- Phillips, J. K., Shafer, J., Ross, K. G., Cox, D. A., & Shadrick, S. B. (2006). *Behaviorally anchored rating scales for the assessment of tactical thinking mental models*. U.S. Army Research Institute for the Behavioral and Social Sciences. Arlington: Department of the Army.
- Pokorny, B. (2012, May). Intelligent tutoring systems: Today and tomorrow. *Military Simulation and Training Magazine*, 21-23.

- Salgado-Zapata, P. J. (1989). *An intelligent computer-assisted instruction system for underway replenishment*. Naval Postgraduate School. Monterey: Naval Postgraduate School.
- Sleeman, D., & Brown, J. S. (Eds.). (1982). *Intelligent tutoring systems*. Orlando, FL, USA: Academic Press, Inc.
- Sullivan, J., & McDowell, P. (2013, January 22). *Learning theories*. Retrieved from Sakai: <https://cle.nps.edu/access/content/group/4acd1024-de79-4f68-b260-d6be2533111c/Week%2003%2017-20%20Jan/Week%203%20Lecture%20Notes/bloomDreyfus.pptx>
- Surface Warfare Officers School Command. (2013, July 22). *(Firearms Training Simulator) System (FATS)*. Retrieved from Surface Warfare Officers School Command: <http://www.netc.navy.mil/centers/swos/Simulators.aspx?ID=5>
- Surface Warfare Officers School Command. (2013, July 22). *Conning Officer Virtual Environment*. Retrieved from Surface Warfare Officers School Command: <http://www.netc.navy.mil/centers/swos/Simulators.aspx?ID=3>
- Surface Warfare Officers School Command. (2013, July 22). *Littoral Combat Ship (LCS)*. Retrieved from Surface Warfare Officers School Command: <http://www.netc.navy.mil/centers/swos/Simulators.aspx?ID=4>
- Surface Warfare Officers School Command. (2013, July 22). *Multi-Mission Tactical Trainer (MMTT)/Tactical Action Officer (TAO) Intelligent Tutoring System (ITS)*. Retrieved from Surface Warfare Officers School Command: <http://www.netc.navy.mil/centers/swos/Simulators.aspx?ID=2>
- Surface Warfare Officers School Command. (2013, July 22). *PC based engineering trainers*. Retrieved from Surface Warfare Officers School Command: <http://www.netc.navy.mil/centers/swos/Simulators.aspx?ID=6>
- Sweller, J. (2011). Cognitive load theory. In B. H. Ross, & J. P. Mestre (Eds.), *The psychology of learning and motivation: cognition in education* (Vol. 55, pp. 37-76). San Diego, CA, USA: Academic Press.
- The Navy Department Library. (n.d.). *Introduction to IFF*. Retrieved May 6, 2014, from The Navy Department Library: <http://www.history.navy.mil/library/online/radar-13.htm>
- United States Government Accountability Office. (2012, June 29). *GAO-12-725R Navy training*. Retrieved from United States Government Accountability Office: <http://www.gao.gov/assets/600/592056.pdf>

- United States Navy. (2013, November 19). *AN/USQT-46 Battle Force Tactical Training (BFTT)*. Retrieved from United States Navy Fact File:
http://www.navy.mil/navydata/fact_display.asp?cid=2100&tid=322&ct=2
- United States Navy. (2013, November 15). *Destroyers - DDG*. Retrieved from United States Navy - Fact File:
http://ipv6.navy.mil/navydata/fact_display.asp?cid=4200&tid=900&ct=4
- University of South Alabama. (n.d.). *Cognitive load theory*. Retrieved May 11, 2014, from University of South Alabama Online Learning Laboratory:
http://www.southalabama.edu/oll/mobile/theory_workbook/cognitive_load_theory.htm
- USC Center for Cognitive Technology. (2013). *Comm Protocol_9_draft*. Redondo Beach, CA.
- Wikipedia*. (2014, April 23). *Unschooling*. Retrieved from Wikipedia:
<http://en.wikipedia.org/wiki/Unschooling>
- Wong, J. H., Kirchenbaum, S. S., & Peters, S. (2010). Developing a cognitive model of expert performance for ship navigation maneuvers in an intelligent tutoring system. *Proceedings of the 19th Conference on Behavior Representation in Modeling and Simulation*, (pp. 29–36). Charleston.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
Ft. Belvoir, Virginia
2. Dudley Knox Library
Naval Postgraduate School
Monterey, California